AGRICULTURAL UNIVERSITY OF ATHENS
DEPARTMENT OF FOOD SCIENCE AND HUMAN NUTRITION
LABORATORY OF MICROBIOLOGY AND BIOTECHNOLOGY OF FOODS

PhD thesis

Characterization of spoilage microbiota through next-generation sequencing and application of spectroscopy-based technologies for assessing the microbiological quality of fresh-cut produce

Evanthia G. Manthou

Supervisor:
George-John E. Nychas, Professor AUA

Supervising Board:
George-John E. Nychas, Professor AUA
Efstathios Panagou, Associate Professor AUA
Panagiotis Skandamis, Professor AUA

ATHENS
2021

AGRICULTURAL UNIVERSITY OF ATHENS
DEPARTMENT OF FOOD SCIENCE AND HUMAN NUTRITION
LABORATORY OF MICROBIOLOGY AND BIOTECHNOLOGY OF FOODS

PhD thesis

Characterization of spoilage microbiota through next-generation sequencing and application of spectroscopy-based technologies for assessing the microbiological quality of fresh-cut produce

«Χαρακτηρισμός των μικροβιακών κοινοτήτων που σχετίζονται με την αλλοίωση μέσω αλληλούχισης επόμενης γενεάς και εφαρμογή τεχνολογιών με βάση την φασματοσκοπία για την αξιολόγηση της μικροβιολογικής ποιότητας φρέσκων κομμένων προϊόντων φυτικής προέλευσης»

Evanthia G. Manthou

Supervising- Evaluation Board
George-John E. Nychas, Professor AUA (supervisor)
Efstathios Panagou, Associate Professor AUA
Panagiotis Skandamis, Professor AUA
Alexandra Lianou**,** Assistant Professor University of Patras
Dimitrios Karpouzas, Professor University of Thessaly
Fady Mohareb, Associate Professor Cranfield University, UK
Stephane Chaillou - Research Director INRAE Institut Micalis, France

Characterization of spoilage microbiota through next-generation sequencing and application of spectroscopy-based technologies for assessing the microbiological quality of fresh-cut produce

Department of Food Science and Human Nutrition
Laboratory of Microbiology and Biotechnology of Foods

Abstract

Fresh-cut produce satisfies the ever-growing consumers' need for healthy and convenient foods. However, the naturally occurring microbial populations along with further microbial contamination during the processing tend to limit the shelf-life of these products. Given that there is limited knowledge of the microbial communities associated with spoilage of these products, the finer characterization of the fresh-cut produce microbiota, and monitoring of the impact of important environmental factors on its composition during storage is needed. On the other hand, the establishment of rapid and effective methods for microbiological spoilage evaluation also is of great importance to the fresh-cut food industry. The two main objectives of the present thesis were: i) to characterize the microbial communities associated with spoilage of these commodities during storage at different temperatures, using a metagenetic approach, (Section A: Chapters 2 and 3), and ii) to evaluate the microbiological spoilage of different fresh-cut produce commodities using rapid spectroscopy-based technologies (Section B: Chapters 4, 5 and 6).

Concerning Chapter 2, the fungal and bacterial communities associated with the spoilage of ready-to-eat (RTE) pineapple during storage at different temperatures were characterized using the ITS2 and *gyrB* metagenetic sequencing, respectively. Significant variability in fungal species composition was observed among the different batches of RTE pineapple. The initial microbiota composition was the main influencing factor determining the progress of spoilage. Depending on the initial prevalent fungal species, the temperature's and storage time's impact varied. With reference to the subdominant bacterial communities, neither temperature nor batch significantly influenced the bacterial diversity and composition.

In Chapter 3, a metagenetic amplicon approach, based on *gyrB* sequencing, was applied for deciphering the bacterial communities associated with the spoilage of RTE rocket and baby spinach stored at different temperature conditions. The vegetable communities differed in the dominance of specific bacterial species. Specifically, *Pseudomonas viridiflava* was dominant

in most samples of rocket, while a new *Pseudomonas* species, as well as *P. fluorescens* and/or *P. fragi* were highly abundant or even dominant in baby spinach. Significant variability in bacterial species composition among the different batches of each vegetable also was observed. At batch-level, the impact of temperature and/or storage time on bacterial microbiota was not apparent for baby spinach. Concerning rocket, the storage time was the most influencing factor for some batches resulting in reduction and increase of *Pseudomonas* species and lactic acid bacteria abundance, respectively. To conclude for both Chapters 2 and 3, a large-scale sampling of fresh-cut RTE produce should be conducted in order to assess the full biodiversity of its spoilage microbiota and unravel the impact of important environmental factors on microbial diversity and composition.

In Chapter 4, the suitability of four analytical technologies (sensors) coupled with different machine learning algorithms for the evaluation of RTE pineapple's quality was assessed, while the utilization potential of selected analytics tools, namely the Unscrambler software and the SorfML platform, also was explored. Pineapple samples stored at different temperature conditions were subjected to microbiological (total viable count, TVC) and sensory analyses, with parallel acquisition of Fourier transform infrared (FTIR), near infrared (NIR), fluorescence (FLUO), visible (VIS), and multispectral imaging (MSI) spectroscopy data. Similar trends concerning the applicability of the different sensors and algorithms were observed for both analytics tools. For TVC, almost all the combinations of sensors and the partial least squares regression (PLSR) algorithm showed relatively satisfactory performances. Moreover, most of the studied sensors in conjunction with linear support vector machine (SVM Linear) resulted in similar performances. Overall, the tested sensors apart from the NIR constitute promising tools for the assessment of pineapple's microbiological spoilage. Concerning sensory features, FLUO spectral data and MSI sensor seem to be also promising for the evaluation of pineapple odour.

Chapter 5 provided a comparative assessment of sensors and machine learning approaches for evaluating the microbiological spoilage of RTE rocket and baby spinach stored at different temperature conditions under modified atmosphere packaging (MAP). The samples were subjected to microbiological analysis (TVC and *Pseudomonas* spp.), and parallel acquisition of FTIR, NIR, VIS and MSI data. Two data partitioning approaches, namely random and dynamic data partitioning, and two machine learning algorithms, namely PLSR

and radial support vector regression (SVR), were comparatively evaluated. Concerning baby spinach stored under passive MAP, the two algorithms yielded similar performances for most of the developed models. The random data partitioning resulted in better (for MSI and NIR) or similar (for FTIR and VIS) performances to the ones attained with the dynamic approach. Regarding rocket, the SVR algorithm resulted in better prediction for most of sensors, while random data partitioning yielded also considerably or slightly better results compared to the dynamic test set. The microbiological spoilage of baby spinach stored under passive MAP was better assessed by models derived mainly from the VIS sensor, while FTIR and MSI models were more suitable in rocket, with the selected best models exhibiting satisfactory performances. Contrarily, FTIR and MSI were more appropriate for TVC prediction of baby spinach stored under active MAP, exhibiting an even higher predictive power. The results indicated that personalised (i.e. product-specific) sensor applications and data analysis workflows are needed, while the applied storage conditions should be also taken into account.

In Chapter 6, the assessment of microbiological spoilage of oyster mushrooms (*Pleurotus ostreatus*) using FTIR and MSI technologies was studied. Fresh-cut mushrooms were stored at different temperature conditions and were subjected to microbiological analysis (TVC) as well as to FTIR and MSI measurements. The FTIR and MSI data were collected for both the cap and the gills sides of mushrooms. Developed PLSR models for both mushroom sides exhibited poor prediction performances, regardless of the applied data partitioning, the reduction of the observed high spectral variability and the selection of the most informative independent variables. The results indicated that, under the conditions of this study and the applied computational analysis, the application of FTIR and MSI do not appear to be promising for the evaluation of the microbiological spoilage of oyster mushrooms.

Scientific area: Food microbiology

Key words: fresh-cut produce, microbiological spoilage, spectroscopy-based sensors, machine learning, microbial diversity, metagenetic analysis, temperature effect

Χαρακτηρισμός των μικροβιακών κοινοτήτων που σχετίζονται με την αλλοίωση μέσω αλληλούχισης επόμενης γενεάς και εφαρμογή τεχνολογιών με βάση τη φασματοσκοπία για την αξιολόγηση της μικροβιολογικής ποιότητας φρέσκων κομμένων προϊόντων φυτικής προέλευσης

Τμήμα Επιστήμης Τροφίμων και Διατροφής Ανθρώπου
Εργαστήριο Μικροβιολογίας και Βιοτεχνολογίας Τροφίμων

Περίληψη

Τα φρέσκα κομμένα προϊόντα ικανοποιούν την συνεχώς αυξανόμενη ανάγκη των καταναλωτών για υγιεινά και ταυτόχρονα γρήγορα και εύκολα στην προετοιμασία τρόφιμα. Βέβαια, οι φυσικά απαντώμενοι μικροβιακοί πληθυσμοί σε συνδυασμό με την επιπρόσθετη επιμόλυνση κατα την διάρκεια της επεξεργασίας τείνουν να περιορίζουν τον χρόνο ζωής αυτών των προϊόντων. Δεδομένης της περιορισμένης γνώσης για τις μικροβιακές κοινότητες που σχετίζονται με την αλλοίωση των φρέσκων κομμένων προϊόντων, ο ευρύτερος χαρακτηρισμός των κοινοτήτων αυτών και η μελέτη των σημαντικών περιβαλλοντικών παραγόντων που επηρεάζουν τη σύστασή τους κατά την διάρκεια της αποθήκευσης των εν λόγω προϊόντων, απαιτείται. Από την άλλη πλευρά, η εφαρμογή ταχέων και αποτελεσματικών μεθόδων για την ανίχνευση της μικροβιολογικής αλλοίωσής τους είναι επίσης σημαντική για την αντίστοιχη βιομηχανία. Οι δύο κύριοι στόχοι της παρούσας διατριβής ήταν: i) ο χαρακτηρισμός των μικροβιακών κοινοτήτων που σχετίζονται με την αλλοίωση διαφορετικών φρέσκων κομμένων προϊόντων κατά την αποθήκευσή τους σε διαφορετικές θερμοκρασίες, με τη χρήση μεταγενετικής προσέγγισης (Ενότητα Α: Κεφάλαια 2 και 3), και ii) η αξιολόγηση της μικροβιολογικής αλλοίωσης των παραπάνω προϊόντων μέσω της χρήσης γρήγορων αισθητήρων με βάση την φασματοσκοπία (Ενότητα Β : Κεφάλαια 4, 5 και 6).

Όσον αφορά στο Κεφάλαιο 2, οι κοινότητες των ζυμών-μυκήτων και των βακτηρίων που σχετίζονται με την αλλοίωση του έτοιμου-προς-κατανάλωση (ready-to-eat, RTE) ανανά κατά την αποθήκευσή του σε διαφορετικές θερμοκρασίες ταυτοποιήθηκαν μέσω της μεταγενετικής αλληλούχισης των γενετικών περιοχών ITS2 και *gyrB*, αντίστοιχα. Παρατηρήθηκε υψηλή παραλλακτικότητα στη σύνθεση των ειδών ζυμών-μυκήτων μεταξύ των διαφόρων παρτίδων ανανά. Η αρχική μικροβιακή σύνθεση ήταν ο κύριος παράγοντας που καθόρισε την πρόοδο της αλλοίωσης. Δεδομένων των αρχικών επικρατέστερων ειδών, η επίδραση της θερμοκρασίας και του χρόνου αποθήκευσης ήταν ποικίλη. Σχετικά με τη μη

κυρίαρχη βακτηριακή κοινότητα, η θερμοκρασία αλλά και η παρτίδα του προϊόντος δεν επηρέασαν σημαντικά την βακτηριακή ποικιλότητα και σύνθεση.

Στο Κεφάλαιο 3, μια μεταγενετική προσέγγιση, βασισμένη στην αλληλούχιση του γενετικού δείκτη *gyrB*, εφαρμόστηκε επίσης για τον χαρακτηρισμό των βακτηριακών κοινοτήτων που σχετίζονται με την αλλοίωση RTE ρόκας και τρυφερών φύλλων σπανακιού, κατά την αποθήκευσή τους σε διαφορετικές συνθήκες θερμοκρασίας. Οι βακτηριακές κοινότητες των δυο λαχανικών διέφεραν στην κυριαρχία συγκεκριμένων βακτηριακών ειδών. Συγκεκριμένα, το *Pseudomonas viridiflava* ήταν κυρίαρχο στα περισσότερα δείγματα ρόκας, ενώ ένα νέο είδος *Pseudomonas*, καθώς και το *P. fluorescens* ή / και *P. fragi* ήταν ιδιαίτερα άφθονα ή ακόμα και κυρίαρχα στο σπανάκι. Παρατηρήθηκε επίσης σημαντική παραλλακτικότητα στη σύνθεση των βακτηριακών ειδών μεταξύ των διαφόρων παρτίδων κάθε είδους λαχανικού. Σε επίπεδο παρτίδας, η επίδραση της θερμοκρασίας ή /και του χρόνου αποθήκευσης στις βακτηριακές κοινότητες του σπανακιού δεν ήταν εμφανής. Σχετικά με την ρόκα, ο χρόνος αποθήκευσης ήταν ο πιο σημαντικός παράγοντας που είχε ως αποτέλεσμα τη μείωση της αφθονίας κάποιων ειδών *Pseudomonas* και την παράλληλη αύξηση της αφθονίας γαλακτικών βακτηρίων. Συμπερασματικά για τα Κεφάλαια 2 και 3, η διεξαγωγή μεγάλης κλίμακας πειραμάτων σε φρέσκα κομμένα προϊόντα φυτικής προέλευσης είναι απαραίτητη για την πλήρη αξιολόγηση της βιοποικιλότητας των μικροβιακών κοινοτήτων που συμμετέχουν στη διαδικασία αλλοίωσης και για την ανάδειξη της επίδρασης σημαντικών περιβαλλοντικών παραγόντων στην μικροβιακή σύνθεση.

Στο Κεφάλαιο 4, η καταλληλότητα τεσσάρων αισθητήρων σε συνδυασμό με διαφορετικούς αλγόριθμους μηχανικής μάθησης αξιολογήθηκε για την εκτίμηση της ποιότητας RTE ανανά, ενώ διερευνήθηκαν επίσης οι δυνατότητες επιλεγμένων αναλυτικών «εργαλείων», του λογισμικού The Unscrambler και της διαδικτυακής πλατφόρμας SorfML. Τα δείγματα ανανά αποθηκεύτηκαν σε διαφορετικές συνθήκες θερμοκρασίας και υποβλήθηκαν σε μικροβιολογική (ολική μεσόφιλη χλωρίδα, OMX) και οργανοληπτική ανάλυση με παράλληλη συλλογή φασματικών δεδομένων φασματοσκοπίας υπέρυθρου με μετασχηματισμό Fourier (FTIR), εγγύς υπέρυθρου (NIR), φθορισμού (FLUO), ορατού (VIS), καθώς και πολυφασματικών εικόνων (MSI). Παρόμοιες τάσεις σχετικά με τη δυνατότητα εφαρμογής των διαφορετικών αισθητήρων και αλγορίθμων παρατηρήθηκαν για τα δυο λογισμικά. Για την OMX, τα μοντέλα από σχεδόν όλους τους συνδυασμούς αισθητήρων με

τον αλγόριθμο παλινδρόμησης μερικών ελαχίστων τετραγώνων (PLSR) παρουσίασαν σχετικά ικανοποιητικές επιδόσεις. Επιπλέον, η γραμμική μηχανή διανυσμάτων υποστήριξης (SVM Linear) σε συνδυασμό με την πλειοψηφία των αισθητήρων παρουσίασε παρόμοιες επιδόσεις. Συνολικά, όλοι οι αισθητήρες εκτός του NIR αποτελούν υποσχόμενα «εργαλεία» για την εκτίμηση της μικροβιολογικής αλλοίωσης φρέσκων κομμένων προϊόντων φυτικής προέλευσης. Σχετικά με τα οργανοληπτικά χαρακτηριστικά, τα FLUO φασματικά δεδομένα και ο MSI αισθητήρας φαίνεται να είναι επίσης κατάλληλοι για την αξιολόγηση της οσμής του ανανά.

Το Κεφάλαιο 5 παρείχε μια συγκριτική αξιολόγηση διαφορετικών αισθητήρων και προσεγγίσεων μηχανικής μάθησης για την εκτίμηση της μικροβιολογικής αλλοίωσης RTE ρόκας και τρυφερών φύλλων (baby) σπανακιού κατά την αποθήκευσή τους σε διαφορετικές θερμοκρασίες και υπό συσκευασία τροποποιημένης ατμόσφαιρας (MAP). Τα δείγματα υποβλήθηκαν σε μικροβιολογικές αναλύσεις (TVC και *Pseudomonas* spp.), με παράλληλη συλλογή φασματικών δεδομένων FTIR, NIR, VIS και MSI. Πραγματοποιήθηκε συγκριτική αξιολόγηση (i) δύο προσεγγίσεων κατανομής των δεδομένων σε σύνολα εκπαίδευσης και επαλήθευσης (τυχαία κατανομή και δυναμικά δεδομένα για επαλήθευση), και (ii) δύο αλγορίθμων μηχανικής μάθησης [PLSR και παλινδρόμηση διανυσμάτων υποστήριξης (SVR) βασισμένησε πυρήνα ακτινικής βάσης (radial basis function, RBF)]. Όσον αφορά στο σπανάκι σε παθητικό MAP, οι δύο αλγόριθμοι παρουσίασαν παρόμοιες επιδόσεις για την πλειονότητα των ανεπτυγμένων μοντέλων. Η τυχαία κατανομή δεδομένων είχε ως αποτέλεσμα καλύτερες (για MSI και NIR) ή παρόμοιες (για FTIR και VIS) επιδόσεις συγκριτικά με τη χρήση των δυναμικών δεδομένων για επαλήθευση. Στην περίπτωση της ρόκας, ο αλγόριθμος SVR οδήγησε σε καλύτερη πρόβλεψη για σχεδόν όλους τους αισθητήρες, ενώ η τυχαία κατανομή δεδομένων έδωσε επίσης σημαντικά ή ελαφρώς καλύτερα αποτελέσματα. Η μικροβιολογική αλλοίωση του σπανακιού σε παθητικό MAP εκτιμήθηκε καλύτερα με μοντέλα που βασίστηκαν κυρίως σε φασματικά δεδομένα VIS, ενώ τα μοντέλα τα οποία αναπτύχθηκαν με βάση δεδομένα FTIR και MSI ήταν πιο κατάλληλα στην ρόκα, όλα με ικανοποιητικές επιδόσεις. Αντίθετα, οι αισθητήρες FTIR και MSI ήταν καταλληλότεροι για την πρόβλεψη της OMX του σπανακιού σε ενεργό MAP, με υψηλότερη ικανότητα πρόβλεψης. Τα αποτελέσματα υπέδειξαν ότι απαιτείται εξατομικευμένη (για κάθε προϊόν) χρήση αισθητήρων και ανάλυση

δεδομένων, ενώ οι συνθήκες αποθήκευσης των προϊόντων θα πρέπει επίσης, να λαμβάνονται υπόψη.

Στο Κεφάλαιο 6, αξιολογήθηκε η εφαρμοσιμότητα των τεχνολογιών FTIR και MSI για την εκτίμηση της μικροβιολογικής αλλοίωσης των μανιταριών στρειδιών (*Pleurotus ostreatus*). Τα φρέσκα μανιτάρια αποθηκεύτηκαν σε διαφορετικές θερμοκρασιακές συνθήκες και υποβλήθηκαν σε μικροβιολογικές αναλύσεις (TVC), με παράλληλη συλλογή δεδομένων φασματοσκοπίας FTIR και MSI και για τις δυο πλευρές (του καπέλου και των ελασμάτων) των μανιταριών. Τα PLSR μοντέλα και για τις δύο πλευρές μανιταριών παρουσίασαν κακή επίδοση πρόβλεψης, ανεξάρτητα από τον εφαρμοζόμενο τρόπο διαχωρισμού των δεδομένων, την προσπάθεια μείωσης της παρατηρούμενης υψηλής φασματικής παραλλακτικότητας που παρατηρήθηκε και την επιλογή των πιο σημαντικών ανεξάρτητων μεταβλητών. Τα αποτελέσματα υπέδειξαν ότι, κάτω από τις συνθήκες της παρούσας μελέτης και της εφαρμοζόμενης υπολογιστικής ανάλυσης, η εφαρμογή των FTIR και MSI δεν είναι υποσχόμενη για την αξιολόγηση της μικροβιολογικής αλλοίωσης των συγκεκριμένων προϊόντων (μανιτάρια στρείδια).

Επιστημονική περιοχή: Μικροβιολογία τροφίμων

Λέξεις κλειδιά: φρέσκα κομμένα προϊόντα, μικροβιολογική αλλοίωση, αισθητήρες με βάση την φασματοσκοπία, μηχανική μάθηση, μικροβιακή ποικιλότητα, μεταγενετική ανάλυση, επίδραση θερμοκρασίας

To my parents,

Giorgos and Areti

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor and supervisor Prof. George – John Nychas for the continuous support, for his motivation, and his guidance all the way through this PhD study. I am also grateful for his trust and for providing me the opportunity to enrich and expand my scientific knowledge.

Besides my supervisor, I would like to thank my co-supervisors: Assoc. Prof. Efstathios Panagou for his insightful comments and advice and Prof. Panagiotis Skandamis for enlightening me the first glance of food microbiology.

I would like to extend my sincere thanks to the rest members of my thesis committee: Prof. Dimitrios Karpouzas, Assoc. Prof. Fady Mohareb, Dr. Stephane Chaillou and Asst. Prof. Alexandra Lianou for dedicating time for my thesis and for their invaluable comments. I would like to offer my special thanks to Dr. Stephane Chaillou who provided me the opportunity to join his team and gave access to the laboratory and research facilities. His continuous support during my stay in France was determinative. My special thanks also go to Asst. Prof. Alexandra Lianou for her treasured scientific and spiritual support. Her immense knowledge and experience have encouraged me and influenced me throughout my academic research.

I would also like to thank Dr. Panagiotis Tsakanikas for his important scientific support and advice, Dr. Agapi Doulgeraki for her valuable advice. Sergio-Llaneza Lago and Gwendoline Coeuret for the excellent collaboration. The members of Institute Micalis at INRAE (National Research Institute for Agriculture, Food and the Environment of France) for embracing me. Of course, big thanks are due to all the members of 'EMBT' (Laboratory of Microbiology and Biotechnology of Foods) for their kind help and the ideal cooperation. Especially, my fellow labmates and also friends Stamatoula Bonatsou, Lemonia Fengou, Eugenia Spyrelli, Eirini Sxoina and Katerina Tzamourani for the continuous support, the stimulating discussions and for all the fun that have made my daily life in and out the lab a wonderful time. My friends Lia, Sofia and Xara for their advice. My friends Cecile, Caroline, Angelo, Marine and many more from Jouy en Josas (France) for a cherished time spent together in France.

Last but not the least, I would like to thank my parents for always believing in me, for encouraging me to follow my dreams and for supporting me spiritually throughout writing this thesis and my life in general.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Chapter 1

1.1. Fresh-cut produce: Current trends

The International Fresh-cut Produce Association (IFPA) defines fresh-cut produce as any fruit or vegetable that has been trimmed, peeled and/or cut into a fully usable product, and is subsequently packaged to offer consumers high nutrition, convenience and flavour, while maintaining freshness. According to the United States Food and Drug Administration (FDA), fresh-cut produce may or may not undergo a wash or other treatment before being packaged, therefore can be ready-to-eat (RTE) or not ready-to-eat (NRTE). However, commodities that have been processed by freezing, cooking, canning or packing in a juice, syrup or dressing, are not included in this category (FDA, 2018). Advances in agronomic, processing, distribution and preservation techniques have enabled the supply of nearly all types of fresh-cut fruits and vegetables to those who desire and are willing to purchase them (Qadri et al., 2015). The current global market includes mainly fresh-cut tropical fruits (melons, cantaloupe, watermelon, mangoes, jackfruit, papaya, grapefruit, pineapples and fruit mixes), shredded leafy vegetables and salad mixes (spinach, lettuce, rocket), fresh-cut vegetables eaten raw or after cooking (peeled baby carrots, baby corn, broccoli and cauliflower florets, cut celery stalks, cut asparagus and cut sweet potatoes), as well as fresh-cut herbs and mushrooms (James & Ngarmsak, 2010; Pradas-Baena et al., 2015).

Fresh-cut products were initially distributed to supply the food service business with the aim to reduce the manpower for food preparation, minimize special systems to handle waste, and deliver specific forms of produce in a short time (Rojas-Graü et al., 2010). Whilst these commodities have been available to consumers since the 1930s in retail supermarkets, it was not before the early 1980s that they started to gradually gain popularity (Baselice et al., 2017). In recent years, the market of fresh-cut produce has witnessed a dramatic growth, largely stimulated by changes in consumers' lifestyle and food needs (Gorni et al., 2015). Fruits and vegetables are a rich source of important nutrients and as such comprise an essential part of balanced diet (Doona et al., 2015). Various international health organizations also recommend the increase of fruit and vegetable intake, since they have been found to counteract many of the chronic diseases, due to compounds with antioxidant and other health-promoting properties (Rojas-Graü et al., 2010; Qadri et al., 2015). Therefore, consumers are increasingly

embracing healthier dietary patterns and have become more concerned about the nutritional and medicinal aspects of the food they consume. At the same time, the modern lifestyle demands easy-to-consume products. Fresh-cut produce is a very convenient way to supply consumers with washed, bite-size and packaged fresh products, thereby RTE commodities that allow to eat healthy on the run save time on food preparation and reduce food waste at the household level (James & Ngarmsak, 2010; Rojas-Graü et al., 2010; Qadri et al., 2015).

Nowadays, the fresh-cut produce supply is one of the growing industries of the agro-food sector, worldwide (Sandhya et al., 2010). It has been converted to a multi-billion-dollar industry, due to the supply of both the food service industry and retail market, as well as its expanding access to new markets across the globe. The main producer and consumer are USA, while UK and France follow (Jideani et al., 2017). According to a market report published in 2010, fresh-cut produce outlined 4% of the total volume of the EU vegetable market and 1% of the EU fruit market. Fresh-cut fruits and salads reached 10 and 50% of the EU fresh-cut markets volume, respectively (Gorni et al., 2015; Qadri et al., 2015; Baselice et al., 2017).

1.1.1. Fresh-cut produce quality and minimal processing

Fresh-cut products rather commonly undergo minimal processing in order to obtain a fully edible product, which is often ready for consumption. The processing steps and their sequence may vary depending on the type of commodity to be minimally processed. While conventional processing methods are known to extend the shelf-life of fruit and vegetables, the minimal processing to which fresh-cut produce are subjected renders products highly perishable, requiring chilled storage to ensure a reasonable shelf-life (Rico et al., 2007). The processing handling including mainly peeling, cutting and shredding steps, accelerates many physiological changes and provides an ideal substrate for growth of microorganisms, including spoilage microorganisms and foodborne pathogens (Rico et al., 2007; Francis et al., 2012).

Fruits and vegetables are living tissues not only when they are attached to the growing parent plant but, even after harvest and until consumption. These products continue to perform most of the metabolic reactions and maintain the physiological systems including respiration, transpiration and ethylene production that were present when they were in their agricultural environment (Kader, 2002; Wills & Golding, 2016). However, the physiology of minimally

processed products is essentially that of living but also wounded tissues. The mechanical damage induced by processing causes stress to plant tissues and accelerates ethylene production, respiration rate, wound-healing processes (synthesis of secondary compounds, suberization and lignification), biochemical (membrane and cell changes, browning and degreening) and physical changes (softening and water loss) (Soliva-Fortuny & Martin-Belloso, 2003; Ergun et al., 2006; Rico et al., 2007). All these biological processes result in flavour loss, off-odours production, cut-surface discolouration, decay, rapid softening, increased rate of vitamin loss and shrinkage, and finally cause irreversible loss of quality (Artes & Allende, 2014; Jideani et al., 2017). A schematic presentation of the primary physiological processes occurring during minimal processing of fresh-cut vegetables resembled with that of fruit is shown in Figure 1.1.

However, microbiological spoilage is an important cause of quality deterioration in fresh-cut produce. These commodities are naturally contaminated by high and diverse microbial populations originating from a number of sources, including the farm environment, pre-harvest treatments, as well as processing where further microbial contamination may occur (Qadri et al., 2015). During processing, the natural protective barriers of cells are removed and the intracellular nutrients and fluids are released. The exposed cytoplasm of plant cells combined with the high water activity (a_w) favor microbial growth (Parish et al., 2003). For these reasons, fresh-cut produce commodities have long been known as vehicles of a large amount of spoilage microorganisms, as well as pathogens (Leff & Fierer, 2013).

In general, microbiological spoilage may manifest as visible growth (slime, colonies), textural breakdown (degradation of polymers), off-odours and off-flavours production (Gram et al., 2002; Ragaert et al., 2007).Various pectinolytic bacteria are able to break down cell walls resulting in exposure of enzymes and substrates, and finally in textural changes and enzymatic browning. The metabolites produced by microbial activity and are related to carbohydrate and protein decomposition, have unpleasant odours and flavours and induce the softening and discoloration of fresh-cut produce. Additionally, vitamins and minerals are utilized by spoilage microorganisms and thereby, pigments and flavour compounds are also degraded, compromising sensory quality (Tournas et al., 2005; Ragaert et al., 2007).

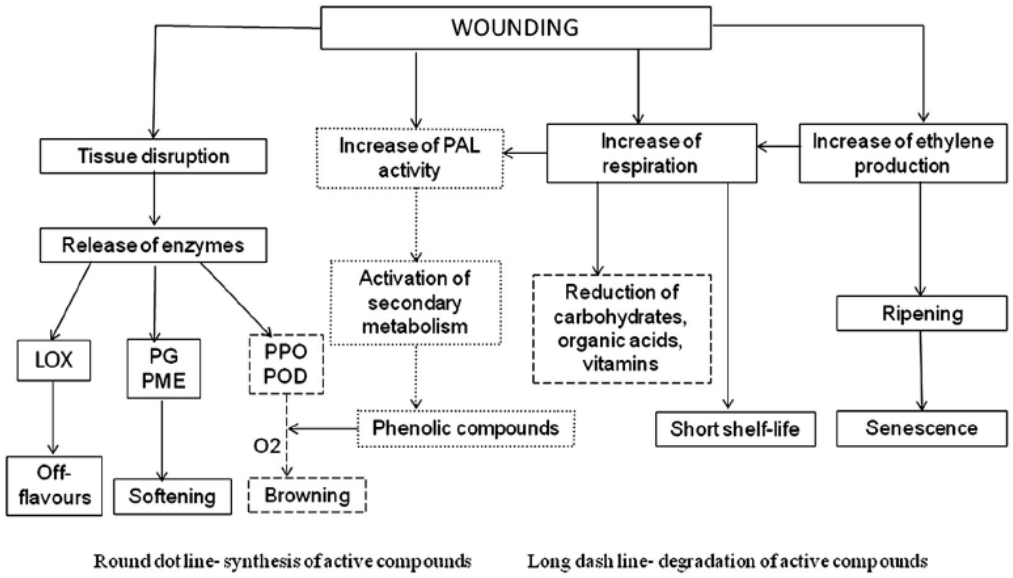


Figure 1.1. Effect of wounding on fresh-cut vegetables. The physiological processes in fresh-cut vegetables after mild processing. Enzymes involved in decay processes. PG: polygalacturonase and PME: pectinmethylesterase cause softening, LOX: lipoxygenase causes off-flavors, PPO: polyphenol oxidase, POD: peroxidase, and PAL: phenylalanine ammonia-lyase, modify the concentration of active compounds during fresh-cut processing. Adopted from Pradas-Baena et al. (2015).

1.1.1.1. Microorganisms associated with fresh-cut produce spoilage

The overall composition, a_w and pH of fresh-cut produce are favourable for microbial growth, including bacteria, molds, and yeasts. Some of these organisms are plant pathogens and they can start the decomposition process before harvest, but most spoilage losses are due to non-plant-pathogens. These organisms can contaminate the commodity in the field, during harvest, transport, processing and/or storage. (Tournas, 2005; Erkmen & Bozoglu 2016).

The predominant post-harvest spoilage of vegetables is caused by Gram-negative bacteria and mainly those capable of degrading the vegetable polymer pectin and causing soft rots. These organisms are typically species belonging to the *Erwinia*, *Pseudomonas* and *Xanthomonas* genera and constitute the specific spoilage organisms (SSO) of several fresh – cut vegetable products (Gram et al., 2002; Qadri et al., 2015). Lactic acid bacteria (LAB) are present in lower populations largerly related to fermentation of vegetables, while yeasts (such

as *Rhodotorula* spp., *Candida* spp.) can also cause spoilage during storage. On the other hand, the low pH of the majority of fruits does not allow growth of most of the bacteria. Molds, yeasts and aciduric bacteria (such as LAB, *Acetobacter* and *Gluconobacter*) are able to grow on fruits (Tournas, 2005, 2006). The yeast genera *Saccharomyces*, *Candida*, *Torulapsis* and *Hansenula* cause spoilage, while *Penicillium*, *Aspergillus*, *Alternaria*, *Botrytis*, *Rhizopus* and other mold genera can cause rot in fresh fruits. However, yeasts are more often associated with spoilage of cut-fruits than molds due to faster growth rate. Lactic and acetic acid bacteria can cause souring in fruits (Erkmen & Bozoglu, 2016).

Moreover, several foodborne pathogens have been reported in fresh-cut vegetables and fruits including the bacteria *Listeria monocytogenes*, *Yersinia enterocolitica*, *Aeromonas*, *Staphylococcus aureus*, *Escherichia coli* O157:H7, *Salmonella enterica* and *Clostridium botulinum* as well as, the fungal species *Aspergillus flavus* and *A. carbonarius* (Ergun et al., 2006; Ragaert et al., 2007; Erkmen & Bozoglu, 2016)

1.1.1.2 Important environmental factors contributing to fresh-cut produce quality

All fresh-cut produce commodities are stored at refrigerated conditions to achieve the aimed commercial shelf-life. Maintaining low temperature after harvest and during processing and storage of fresh-cut produce in retail market is the most critical aspect of fresh-cut produce quality. Temperature has a great impact on metabolic reactions and it strongly affects respiration rate, ethylene production and changes in atmosphere within commercial packages (Kader, 2002; Francis et al., 2012). Moreover, storage temperature is one of the most important factors affecting survival and growth of spoilage microorganisms and pathogens on fresh-cut produce (Gu et al., 2018). Hence, effective cold chain temperature control is of great importance for fresh produce, since often there is no thermal treatment before consumption (Castro-Ibanez et al., 2017)

Modified atmosphere packaging (MAP) has also the potential to extend the shelf-life of fresh-cut produce, mainly by limiting the oxidation processes and the proliferation of aerobic spoilage microorganisms reaching specific levels of O_2 and CO_2 (Artes & Allende, 2014). This type of packaging relies on the rate of respiration of produce together with the transfer of gases through the packaging material without any further alteration to the initial composition. Passive MAP involves generation of modified atmosphere by reliance wholly on the natural

process of respiration as well as the, permeability of the packaging film material. On the other hand, active MAP involves the replacement of the gaseous composition in a packaged material establishing the preferred gas mixture within the package (Jideani et al., 2017). The beneficial effects of MAP include the reduction of respiration rate, ethylene production, enzymatic reactions, and some physiological disorders (Sandhya et al., 2010; Francis et al., 2012). However, the overall effects of modified atmospheres (MA) have been extensively evaluated on a wide range of fruit and vegetables, but the responses vary considerably (Wills & Golding, 2016). The limits of tolerance to O_2 and CO_2 levels (before resulting in physiological damage with subsequent effects on sensory quality) depend on type of produce, cultivars, temperature, physiological condition, maturity and previous treatment (Francis et al., 2012; Ma et al., 2017). With regards to microbial activity, the effect of MA is not consistent and the storage temperature tends to largely control microbial growth (Ragaert et al., 2007). The application of high O_2 or superatmospheric levels have been reported to inhibit aerobic and anaerobic microbial growth. However, CO_2 is the only gas that has a direct and significant antimicrobial effect, but its inhibitory effect is not universal and is dependent on the microorganism and its growth phase, the temperature, a_w and product characteristics (Oliveira et al., 2015). On the other hand, excessively low levels of O_2 favour fermentative processes which might cause the formation of acetaldehyde and the appearance of off-flavour compounds (Artes & Allende, 2014).

1.2. Advanced molecular technologies for the study of food microbial ecology

Deep knowledge and understanding of microbial food ecology is crucial to ensure production of safe and high-quality food. In particular, close monitoring of the diversity and dynamics of food microbial communities throughout the production process allows the effective management of the microbial processes involved in food processing, spoilage and contamination (Juste et al., 2008; Cao et al., 2017).

Traditionally, the occurrence of microorganisms in a given environment or in an industrial process has been studied by culture-based methods. So far, culture-dependent methods have been the mainstay of food microbiology, since they have led to the description of a number of habitats. However, they are extremely biased in their ability to unravel the

microbial communities of complex matrices associated with food or environmental samples (Ercolini et al., 2013; Zhou et al., 2015). Indeed, they are able to detect only 0.1% of a complex community and even in relatively simple food matrices, such as fermented foods, at least 25–50% of the active microbial community cannot be cultured *in vitro*. Moreover, stressed or weakened cells often need specific culture conditions to recover and to become cultivable, while there are also entirely non culturable cells (Juste et al., 2008).

These limitations have prompted the development of culture-independent techniques, of which those based on polymerase chain reaction (PCR) amplification and detection of nucleic acids are the predominant (Edet et al., 2017). Nucleic acids make the molecule an ideal target for microbiota or microbiome characterization due to their ubiquitous nature and specificity. Additionally, significant progress in sequencing chemistries have resulted in evolution from low throughput sequencing to high throughput next generation sequencing (NGS) technologies providing a broader microbial characterization (Cao et al., 2017).

1.2.1. Next generation sequencing technologies

The development of NGS technologies has enabled researchers to study food microbial ecology from broader and deeper perspectives, revealing that food-associated microbial communities are richer than expected and that some of them might play a yet unsuspected role (Kergurlay et al., 2015). In the case of microbial spoilage, NGS has opened up new perspectives of characterization and control management, since spoilage microbiota for most of food products had remained poorly characterized due to their diversity and the lack of selective culture media for their isolation. Such approaches can readily yield multiple thousands of sequences from a single sample providing much more coverage of the community present, and increasing the ability to detect less common taxa. Moreover, NGS offers the opportunity to look beyond the presence of taxonomically defined entities and instead to understand the relationships between microorganisms as well as their activities and functionalities in a particular niche (Jackson et al., 2015; Cocolin et al., 2018).

Currently, the majority of microbial ecology studies apply NGS by focusing on either targeted gene sequencing with phylogenetic (e.g., 16S rRNA) and also functional (e.g., *amoA*, *nifH*) gene targets or on shotgun metagenome sequencing. Targeted gene sequencing (metataxonomics, metagenetics, metabarcoding or amplicon sequencing) provides information

on specific genes within a microbial community, and is thereby a “taxonomy” oriented approach. However, its suitability for analyzing the whole genetic and functional diversity of communities is limited. On the other hand, metagenomics is an approach that is mainly “function” oriented, although taxonomic composition of communities can also be revealed. Metagenomics focus on the presence of genes and their transcripts, rather than the identification of a specific organism (Cocolin et al., 2018). However, both metagenetics and metagenomics cannot distinguish between expressed and non-expressed genes in a given environment. In contrast, metatranscriptomics involve random sequencing of expressed microbial RNA. In Figure 1.2, the most common NGS applications in applied microbiology are reported. However, further development of proteomics and metabolomics is important to understand microbial community functions in the environment (Zhou et al., 2015; Cao et al., 2017).

1.2.1.1. Targeted gene sequencing (metataxonomics, metagenetics, metabarcoding or amplicon sequencing)

The use of targeted gene sequencing to study microbiota is the most common NGS application in microbial ecology of foods (Ercolini et al., 2013). The terms metataxonomics, metagenetics, metabarcoding or amplicon sequencing refer to the amplification and high-throughput sequencing of a specific barcode region (e.g. 16S for bacteria, 18S for most eukaryotes, and ITS regions for fungi). This approach enables the study of all the microbial taxa in a given environment, which is defined as the microbiota. It is important to underline that microbiota refers to the taxonomic composition of the microbial community as determined by metagenetic analysis, while microbiome refers to the entirety of the microbial genetic material determined by metagenomics (Ursell et al., 2012).

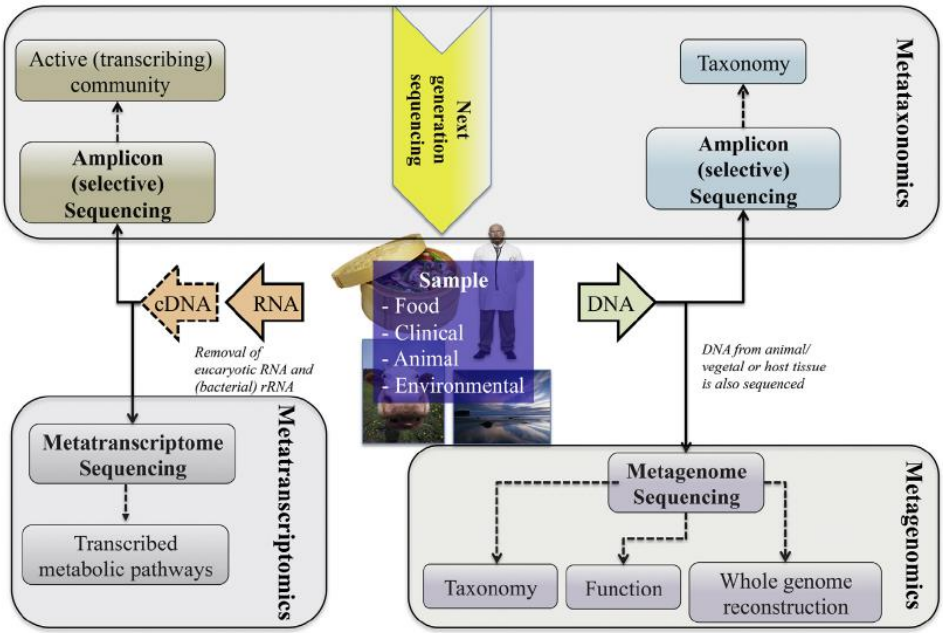


Figure 1.2. The use of next generation sequencing (NGS) in applied microbiology can be used to better understand ecology and interactions of microorganisms in specific complex ecosystems. Not only diversity, but also behaviour can be investigated through targeting of DNA and RNA, respectively. Adopted from Cocollin et al. (2018).

The metagenetic approach includes five main steps: i) sampling; ii) DNA extraction; iii) targeted amplification; iv) sequencing and v) data analysis. A more detailed workflow is presented in Figure 1.3 (Abdelfattah et al., 2018). Roughly, DNA is extracted from food samples using various extraction and purification methods. Specific conserved primers are used for targeted genes amplification which are barcoded with short oligonucleotide tags and sequencing adapters, so that multiple samples can be pooled and sequenced simultaneously. Then, target DNA is quantified, sequenced, and finally analyzed using advanced bioinformatics tools (Zhou et al., 2015).

The sequencing step can be performed on different platforms that have been developed and are widely used for NGS applications. Since multiple NGS platforms are available in the market, the decision has to be made depending on the desirable outcome and taking into account the varying features of platforms such as read length, degree of automation, throughput per run, data quality, ease in data analysis and cost per run (Kumar et al., 2015). Illumina® and Thermo Fisher Scientific® have produced sequencers that are the most commonly used in

metabarcode sequencing of microbial communities. These are the MiSeq™, NextSeq™ and HiSeq™ from Illumina® (San Diego, California) and IonTorrent™ and IonProton™ from Thermo Fisher Scientific® (Waltham, Massachusetts) (Zhou et al., 2015; Abdelfattah et al., 2018). Illumina platforms are based on sequencing by synthesis method using cyclic reversible termination (CRT). Thise approach is defined by the terminator molecules in which the ribose 3´ OH group is blocked, preventing elongation. To begin the sequencing, DNA template is primed by a sequence that is complementary to an adapter region, which will initiate polymerase binding to this double-stranded DNA. For dNTP identification, total internal reflection fluorescence (TIRF) microscopy using either two or four laser channels is applied, since each dNTP is bound to a single fluorophore that is specific to that base type. Illumina CRT system accounts for the largest market share for sequencing instruments compared to other platforms (Goodwin et al., 2017).

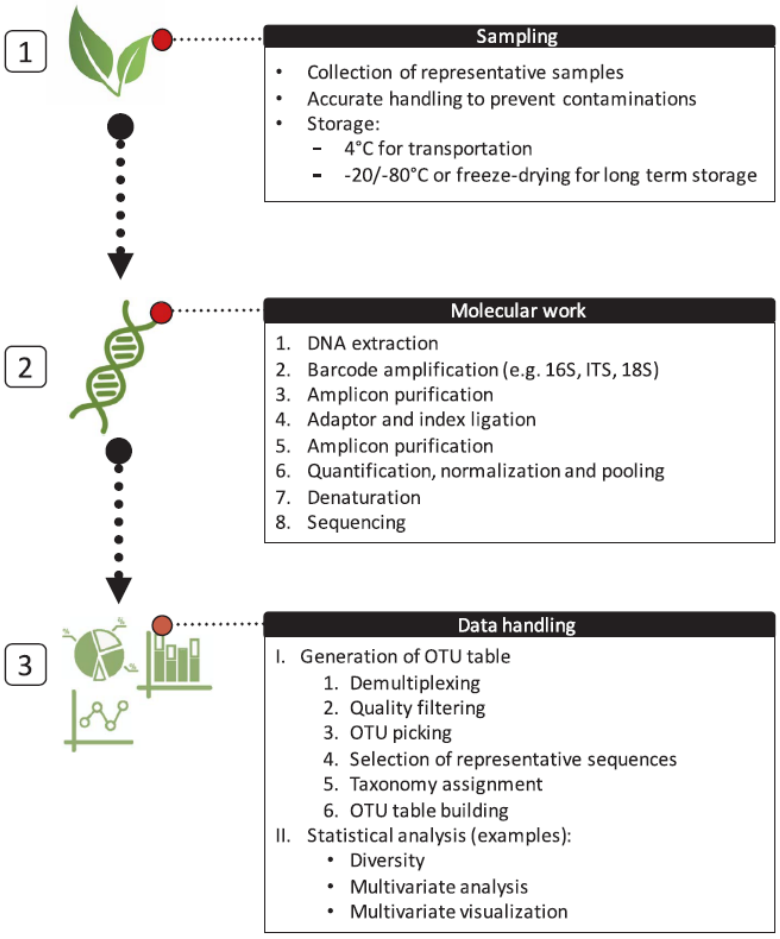


Figure 1.3. An outline of the metabarcoding approach. Adopted from Abdelfattah et al. (2018).

The analysis of the mass high throughput sequencing data derived from these platforms requires advanced bioinformatics tools before the results can be interpreted (Zhou et al., 2015). The development of various open-source data storage and bioinformatics pipelines such as QIIME, MOTHUR and Galaxy web server (Schloss et al., 2009; Caporaso et al., 2010; Afgan et al., 2018) for analyzing community sequence data, has enabled researchers with limited expertise in bioinformatics to undertake elaborative projects in meta-omics. These pipelines involve various processing steps for the final generation of an Operational Taxonomic Units (OTUs) table (Figure 1.4). First, the demultiplexing step assigns reads to each biological sample according to the barcodes added during the PCR. The quality filtering is a step of great importance aiming to discard reads with low quality base calls, chimeras and short reads that could lead to erroneous results. Once the sequences are filtered, they are clustered using algorithms that rely on the selection of sequences containing similar sequences according to a pre-selected threshold of similarity. After this process, a representative set of sequences is generated, where one representative sequence corresponds to each cluster. The representative OTU sequence is then queried against a set of established databases for taxonomic assignment, such as RDP for 16S, SILVA for 18S and UNITE for ITS2 (Wang et al., 2007; Quast et al., 2013; Nilsson et al., 2019), or custom databases. Once the OTUs table is constructed, a wide range of statistical analyses can be carried out between the samples. Before further analyses, the OTUs table should be normalized due to variations in the size (number of reads) of each sample. The main analysis of the OTU data commonly includes the characterization of alpha and beta diversity, and statistical comparisons between samples with a wide set of multivariate statistical procedures (e.g. ANOVA) and visualization tools (e.g. Principal Coordinates Analysis (PCoA), Non-metric multidimensional scaling (NMDS) and heatmaps (Abdelfattah et al., 2018).

1.2.1.2. Genetic markers

The bacterial 16S and the fungal ITS genes are the most widely used target barcodes for studying the phylogenetic relationships present into the food-associated communities. Several studies focus on the selection of specific sub-regions such as hypervariable regions from V1-V9 in bacteria and ITS1 or ITS2 for fungi for improved characterization results. Both genetic markers are considered ideal candidates for revealing microbial diversity since they

consist of conserved regions, which serve as annealing sites for the design of universal PCR primers. Their extensive utilization has also led to the development of large sequence databases facilitating the analysis. However, taxonomic methods based on 16S rDNA and ITS profiling analyses have also important shortcomings (Abdelfattah et al., 2018).

The extremely slow rate of evolution of 16S rDNA gene hinders the resolution of closely related bacteria into individual phylotypes. Regarding the regions used, the obtained results may vary widely. Many authors have favoured different hypervariable regions. The rDNA V1-V3 region has been used in many studies of food microbiota due to the sufficient discrimination among species of lactic acid bacteria. However, the requirement of short reads for the recent and more accurate technologies (such as Miseq pair-end sequencing) has shifted the focus on the rDNA V3-V4 region. This region does not exhibit the same discriminatory power as the V1-V3 region for identifying species-level diversity and various difficulties in taxon quantification are encountered. Consequently, 16S rDNA amplicon data are often analyzed at the genus level only, lacking the power to yield informative answers to food spoilage and food processing conditions which can be species- or even strain-specific (Cao et al., 2017; Poirier et al., 2018). Similarly, insufficient genetic variability within the ITS regions has been also acknowledged. The limited variability present in this barcode gene can make the discrimination of closely related taxa very difficult, mainly for the phylogenetic studies aimed at redefining the classification of fungi (Abdelfattah et al., 2018).

The genes that have been proposed instead of /together with 16S include those encoding 23S rRNA, DNA gyrase subunit B (*gyrB*), RNA polymerase subunit B (*rpoB*), TU elongation factor (*tuf*), DNA recombinase protein (*recA*), protein synthesis elongation factor-G (*fusA*), and dinitrogenase protein subunit D (*nifD*). Among these, the *gyrB* gene has a higher rate of base substitution than 16S rDNA, and shows promising results for community-profiling applications. This gene encodes the subunit B of DNA gyrase, a type II DNA topoisomerase which catalyses breakage and reformation of double-stranded DNA. It is a single-copy gene and is sufficiently large in size for use in analysis of microbial communities (Cocolin et al., 2013; Leo et al., 2013; Poirier et al., 2018). To date, *gyrB* has been sparingly used in food microbiota studies with promising results at species-level community characterization (Poirier et al., 2018, 2020; Manthou et al., 2020; Zagdoun et al., 2020). In the case of fungal characterization, the translation elongation factor 1-α (TEF1α), topoisomerase I (TOPI),

phosphoglycerate kinase (PGK) and the RAS-related YPT1 gene have been also proposed as effective universal fungal DNA barcodes (Abdelfattah et al., 2018).

1.3. Rapid technologies for detection of microbiological spoilage

Microbiological spoilage plays an important role in the limited shelf-life of fresh-cut commodities, affecting significantly their quality and resulting in remarkably high economic loss for the fresh-cut industry (Granato et al., 2018; Giannoglou et al., 2020). Therefore, apart from the development of strategies for the suppression of microbial risks during production and storage, the establishment of effective and rapid detection methods of food spoilage are of great importance for the fresh-cut industry (Wang et al., 2018).

Conventional microbiological methods, including culturing and colony-counting methods as well as biochemical assays are the gold standards in food microbiology. They are cost-effective, easy to-use, and show limited requirements of specialized instruments (Ripolles-Avila et al., 2020). Molecular based techniques are also widely used in food microbiology as more reliable, accurate and rapid tools, even though high-tech instruments and trained personnel are required. Nevertheless, all these methods are time-consuming providing retrospective results, and usually destructive to test products, limiting thus their potential to be used on-, in- or at-line (Nychas et al., 2016).

Compared to the aforementioned methods, spectroscopy and spectral imaging approaches have become popular and attractive due to minimal sample preparation, non-destructive sampling, rapid data acquisition, as well as detection during the production line (Nychas et al., 2016; Efenberger-Szmechtyk et al., 2018) (Figure 1.4). These methods have been successfully reported in the literature as promising tools for quality, safety and/or authentication-adulteration assessment of different meat products, such as poultry (Barbin et al., 2015; Grewal et al., 2015; Spyrelli et al. 2020), pork and beef (Argyri et al., 2013; Oto et al., 2013; Ropodi et al., 2015; Trinderup et al., 2015; Fengou et al., 2019a), as well as fish (He & Sun, 2013; Duan et al., 2014; He & Sun, 2015a; Saraiva et al., 2017; Fengou et al., 2019b). Their application to fruits (Camps & Christen, 2009; Unay et al., 2011; Liu et al., 2015; Al-Holy et al., 2015; Manthou et al., 2020) and vegetables (Wang et al., 2010; Løkke et al., 2013; Sravan-Kumar et al., 2015; Tsakanikas et al., 2018; Bureau et al., 2019) but also olive oil (de la Mata et al., 2012; Guzmán et al., 2015) and dairy products (Nicolaou & Goodacre, 2008;

Subramanian et al., 2011; Jacquot et al., 2015; Lianou et al., 2019), has also been demonstrated. The qualitative or quantitative evaluation of microbiological spoilage has been reported mainly for meat and seafood products, but limited is the scientific work related to the evaluation of microbiological spoilage of fruits and vegetables.

1.3.1. Spectroscopy and multispectral imaging technologies

The analytical technologies based on vibrational (infrared and Raman) spectroscopy, fluorescence and ultraviolet-visible spectroscopy, as well as surface chemistry (hyper/multispectral imaging) have been widely used in food quality assessment (Lohumi et al., 2015). Recently, the aforementioned technologies have been also utilized for the rapid and quantitative monitoring of microbiological spoilage. The general principle of these approaches is that by-products of microbial metabolic activity display different biochemical profiles resulting in a characteristic sensor fingerprint that could be used for quality evaluation (Ellis & Goodacre, 2001; Nychas et al., 2008). The technologies applied in the experimental procedures of the present thesis are described in more detail in the following sub-sections.

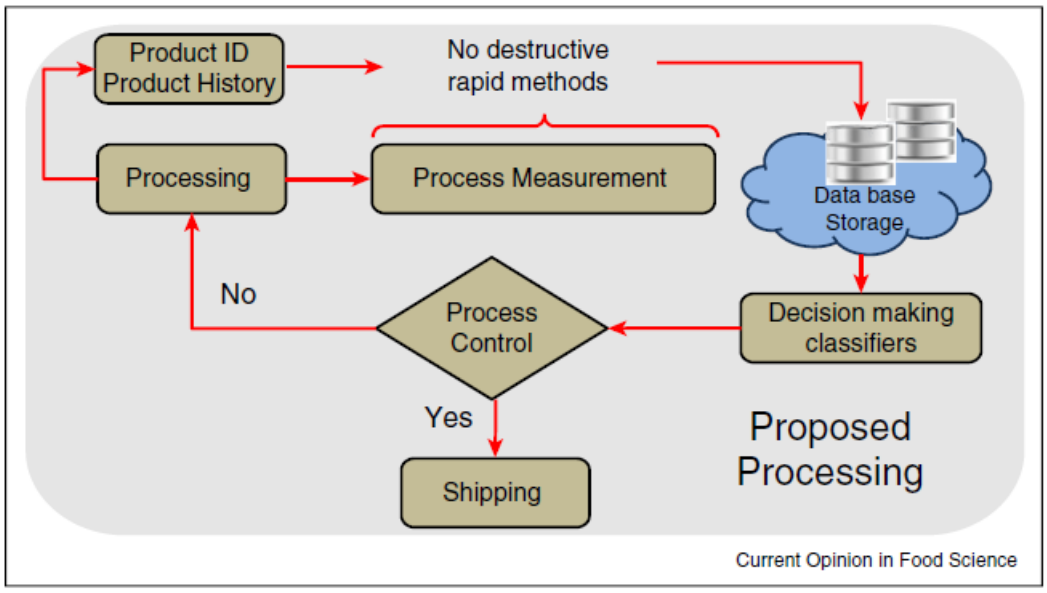


Figure 1.4. Use of non-destructive rapid methods in the food industry. Adopted from Nychas et al. (2016).

1.3.1.1. Infrared spectroscopy (IRS)

Infrared (IR) spectroscopy is a technique based on the vibrations of the atoms of a molecule. The energy at which any peak in an absorption spectrum appears corresponds to the frequency of a vibration of a part of a sample molecule. The IR region can be subdivided into near-infrared (12500 to 4000 cm^{-1}), mid-infrared (4000 to 400 cm^{-1}) and far-infrared (400 to 33 cm^{-1}) regions. Both the near-infrared (NIR) and mid-infrared (MIR) regions can be used for studying the quality of food products, since the absorption in these spectral ranges can be related, to a greater or lesser degree, to the main chemical components of foods such as proteins, carbohydrates, fats and water. They have been also employed for the detection of microorganisms. Microorganisms have unique chemical components in their cell membrane and cell wall, which in turn can provide distinct IR absorption spectra (Di Egidio et al., 2009; Wang et al., 2018).

The MIR region provides information on fundamental frequencies of chemical bonds in functional groups such as C–C, C–H, O–H, C= O and N–H. A typical MIR spectrum may be roughly divided in two regions, the functional group region from 4000 to 1500 cm^{-1} and the fingerprint region from 1500 to 500 cm^{-1}. The fingerprint region is normally complex with many overlapped bands, but it is also specific to the molecular structure of samples (Di Egidio et al., 2009; Lohumi et al., 2015; Zhang et al., 2018). There are many absorption bands in the MIR region arising from various functional groups present in water, lipids, proteins, polysaccharides and nucleic acids. The absorption peaks around 3400 cm^{-1} are primarily from water. The absorption peaks around 2960, 2929, and 1740 cm^{-1} are from fatty acids, while peaks around 1650 and 1550 cm^{-1} correspond to amide I and II vibrations of protein or peptides. In addition, the peaks between 1200 and 900 cm^{-1} originate from the stretching vibrations of the phosphate and the vibrations of polysaccharide compounds (Lin et al., 2009).

The NIR region is associated with molecular vibrations, specifically the overtones (secondary vibrations) and combinations of fundamental vibrations. Chemical bonds between light atoms, such as O-H, N-H C-H, and S-H have high vibrational frequencies, which result in overtone and combination bands detectable in this region. These bands are mostly affected by the formation of hydrogen bonds and can generate a characteristic spectrum, thereby the fingerprint of the sample for chemical structure monitoring (Lohumi et al., 2015; Wang et al., 2018).

1.3.1.2 Ultraviolet-visible spectroscopy and fluorescence spectroscopy

Ultraviolet-visible (UV-VIS) absorption spectroscopy and fluorescence (FLUO) emission spectroscopy correspond to the same range of wavelengths, but are caused by two different phenomena. In both cases, the wavelengths used are the near-ultraviolet range (200 to 400 nm) and the visible range (400 to 750 nm). However, UV-VIS measures the absorption of light in these two regions, while fluorescence measures the light emitted by a sample in this range after absorbing light at a higher energy than it is emitting.

The phenomenon of UV and visible light absorption is restricted to specific chromophores and several chemical species with defined molecular functional groups. In particular, UV-VIS spectroscopy probes electronic excitations of molecules that are typically related to absorption of photon energies corresponding to the two spectral regions (Power et al., 2019). On the other hand, some molecules, promoted to an electronically excited state by absorption of UV or visible radiation, return to their ground state by emission of photons with an energy different from that absorbed. Therefore, fluorescence is the emission of light by a fluorophore following the absorption of UV or VIS light. Fluorophores are fluorescent molecules characterized by conjugated systems of double C=C bonds, aromatic character and rigid molecular skeletons, and include polyaromatic hydrocarbons, heterocyclic compounds, and a few highly unsaturated aliphatic compounds (Hassoun et al., 2019; Sikorska et al., 2019).

1.3.1.3 Multispectral imaging

Multispectral imaging (MSI) is a spectral imaging technology based on hyperspectral imaging (HSI). Both of them integrate both imaging and spectroscopic units into one system and provide a 3D hypercube containing rich spectral as well as spatial information. The distinctive difference between these two technologies depend mainly on the number of spectral bands involved in the hypercube. For HSI, there are more than 100 contiguous and regular spectral bands, whereas normally 20 non-contiguous and irregularly spectral bands are applied in the case of MSI. Although MSI systems cannot provide fine details in the spectral signatures in every image pixel, the instrumental complexity and cost, as well as the data acquisition time, are significantly lower compared to those of HSI systems (Zhang et al., 2018). Various instruments have been developed to provide information over the UV, VIS and NIR regions of the electromagnetic spectrum.

1.3.2. Data analysis

Despite the advantages of spectroscopy and imaging techniques, mass data are generated by these analytical technologies posing an extreme challenge for data mining and processing (Zhou et al., 2019). The multivariate nature of the analytical output is rather complex and its analysis demands a multi-disciplinary approach. Depending on the type of sensor and data complexity, the analysis may involve computer vision / image processing, signal processing, statistical analysis, machine learning and/or other advanced computational techniques (Jollife & Cadima, 2016; Ropodi et al., 2016; Truong et al., 2019).

In most applications of analytical technologies, the main objective is the creation of a model that correlates dependent and independent variables and predicts the unknown. Machine learning approaches provide related information through the development of classification or regression models using spectral or imaging data for model training and validation (Nychas et al., 2016; Ropodi et al., 2016). Advances in data science have introduced a plethora of machine learning approaches applied in tandem with the above analytical techniques. However, choosing the appropriate machine learning approach based on the question that should be addressed, is often challenging and involves a comparative analysis in order to achieve the best possible solution (Estelles-Lopez et al., 2017).

1.3.2.1. Data pre-processing and pre-treatment

Prior to the specific statistical analysis, the data are often subjected to different pre-processing and pre-treatment steps. According to Goodacre (2007), pre-processing is a generic term for methods to go from raw instrumental data to clean data for data processing, while pre-treatment is the transformation of the clean data to make them ready for data processing. All these steps allow the signal from dissimilar samples to be compared as well as to eliminate effects of unwanted signals such as fluorescence, scattering effects, detector noise, calibration errors, cosmic rays, laser power fluctuations, signals from cell media or glass substrate (Gautam et al., 2015). The selection of the most appropriate method is highly dependent on the sample, instrumentation technique and purpose of analysis. The most widely used pre-processing and/or pre-treatment methods in spectroscopy data are normalization, standardization, multiplicative scatter correction (MSC), standard normal variate (SNV),

Savitzky-Golay algorithm and Savitzky-Golay polynomial derivative filters (Rinnan et al., 2009; Gautam et al., 2015; Roberts & Cozzolino, 2016).

1.3.2.2. Dimensionality reduction

Spectroscopy techniques have the capability to generate very large amounts of data. The simplest decision is to include all of the available features in further data analysis reducing the possibility to lose any important information. On the other hand, the irrelevant variables may negatively influence the estimation/predictive performance of the developed statistical models. Moreover, machine learning suffers from the phenomenon of dimensionality, which is related to the existence of many variables and few samples. In particular, learning a model in high dimensions is more difficult than learning a model in lower dimensional spaces. As a result, some machine learning methods perform multivariate dimension reduction as a part of the learning process (Mehmood et al., 2012; Torrione et al., 2014).

The data reduction approaches may be performed through projection methods, variable selection or a combination of both. A wide range of statistical techniques have been applied in research studies including principal components analysis (PCA) and partial least squares (PLS), as the most popular in spectroscopy data, as well as the most recently introduced random forests (RFs) (Torrione et al., 2014; Hu et al., 2018; Tsakanikas et al., 2018, 2020). However, it should be emphasized that variable selection may also eliminate some useful redundancy from the model, and using a small number of variables for prediction means large influence of each variable in the final model (Mehmood et al., 2012).

1.3.2.3. Machine learning algorithms and model development

Machine learning methods can be divided into: supervised and unsupervised pattern recognition. Unsupervised pattern recognition reveals clusters according to the sample similarity without prior knowledge and assumptions. Among the unsupervised methods, hierarchical cluster analysis (HCA) and PCA have been extensively used in food applications (Argyri et al., 2014). In supervised pattern recognition, the developed model is trained using an input learning subset, in order to unravel hidden patterns within the data and predict a target variable or class. The prediction can be either a nominal value (classification model) or a numeric value (regression model). These techniques include partial least squares regression

(PLSR), partial least squares discriminant analysis (PLSDA), k–nearest neighbour (kNN) and support vector machines (SVM) algorithms (Efenberger-Szmechtyk et al., 2018). A more detailed description of supervised pattern recognition methods is provided below.

The **PLS/PLSDA** methods belong to linear regression methods that are extensively used in food applications, since they are relatively easy to apply. The PLSR method projects both observed and predicted values in a feature space and a linear regression model is established. It can be also used for classification purposes when coupled with linear discriminant analysis (LDA) (Barker & Rayens, 2003). The PLSR and PLSDA methods are appropriate for spectroscopic data, since they can deal with high data dimensionality (where the generated data contain many variables and insufficient number of samples) as well as with strong collinearity and noise (Wold et al., 2001; Mehmood et al., 2012).

The **kNN** is a simple method that does not use any type of distribution assumption and can be used for a small number of samples. This method can be used for classification of categorical data variables and regression for continuous variables (Silverman & Jones, 1989; Granato et al., 2018). In particular, it uses a similarity measure for comparing the testing data with training data. For prediction of output variables, it chooses k data points from the training dataset that are close to the testing dataset. Therefore, this method does not build a model or a function, but yields the closest k records of the training dataset that are the most similar to the points that are to be categorized or predicted (Al-Dosary et al., 2019).

The **SVM** is an effective machine learning technique suitable for both classification and regression problems. This method is based on mapping data into a high-dimensional space that allows for the separation of two groups of samples into distinctive regions using a kernel function, in order to construct a maximal separating hyper-plane (Cortes & Vapnik, 1995). This separation is achieved by identifying only a small fraction of the samples, referred to as 'support vectors', between which the separating hyper-plane is identified (Gromski et al., 2018). There are various choices for the kernel functions applied, including linear, polynomial, sigmoid and radial basis function (Ropodi et al., 2016).

The **RFs** method generates many decision trees (classification and regression trees, CART), each of which is constructed using different sets of randomly selected input (X) variables (Breiman, 2001). The trees are created by drawing a subset of training samples through replacement (bootstrapping). This means that the same sample can be selected several

times, while others may not be selected at all. The data are segmented in a series of training sets, which on average include 63.2% of all samples, and a series of test sets, which includes the remaining 36.8% samples. The training datasets are used to construct the trees whereas the test sets are used to estimate model performance (Belgiu & Dragut, 2016; Gromski et al., 2018). Recently, a fast version of RFs, namely Ranger, has been introduced and is appropriate for high dimensional data (Wright & Ziegler, 2017).

1.3.2.4. Model validation

After the calibration model has been developed, validation is an integral part of the process in order to test the predictive ability of the model developed. The model capability to predict unknown samples must be demonstrated using an independent set of samples. It is generally agreed that the independent sets of samples should be sourced from other experiments, batches or conditions. Model validation in its simplest form involves splitting the data into training and test sets. The training data are used to build one or more possible models and the independent test data are used to measure the generalization / predictive performance of the model. The data partitioning may be performed empirically or through different methods or algorithms such as PCA, Bayesian, Kennard–Stone algorithms and random sampling.

Internal validation (cross-validation) is used to optimize meta-parameters (e.g., number of latent variables in PLS), but it is not a replacement for external model validation. There are several methods in which the training and cross-validation sets are initially combined and then subsequently partitioned into temporary training/cross-validation datasets, including bootstrapping, leave-one-out (LOOCV), K-fold cross-validation (Goodacre et al., 2009; Roberts & Cozzolino, 2016; Truong et al., 2019).

1.3.2.5. Model performance

Numerous statistics have been described to interpret the results obtained during model development. In the case of regression models, mainly the root mean squared error (RMSE), but also the standard error of prediction (SEP) or the residual predictive deviation (RPD) value are used as performance indices and provide an estimation of the average uncertainty that can be expected for predictions of future samples. Another commonly used statistic is the coefficient of determination (R^2), which represents the proportion of explained variance of the

response variable in either the training or test sets (Truong et al., 2019). Additionally, the correlation coefficient (r), slope and offset of the linear regression between the predicted and actual (measured) dependent variables are also used for evaluating the ability of a calibration model to predict new samples (Papadopoulou et al., 2011; Tsakanikas et al., 2018). In the case of classification, the values of specificity, sensitivity, precision and accuracy are indicative of the model performance (Ropodi et al., 2016).

1.4. Objectives of the thesis

The present thesis consists of two main sections, Section A and Section B. The former includes the Chapters 2, 3 and concerns the metagenetic characterization of the microbial communities associated with the spoilage of fresh-cut produce commodities during their storage under different temperature conditions. The latter section includes the Chapters 4, 5 and 6 concerning the assessment of the microbiological quality of fresh-cut produce through rapid, non-invasive spectroscopy-based technologies coupled with machine learning.

In Chapter 2, the fungal and bacterial communities associated with the spoilage of RTE pineapple were characterized through the metagenetic sequencing of the ITS2 and *gyrB* genetic regions, respectively. Moreover, the variability of RTE pineapple's spoilage microbiota as affected by temperature during storage was also assessed.

Chapter 3 describes the bacterial communities related to RTE baby spinach and rocket spoilage. Additionally, the changes occurring in their composition and diversity during storage at different temperatures was also evaluated. The bacterial characterization was performed through the metagenetic amplicon approach based on *gyrB* sequencing.

The objective of **Chapter 4** is the assessment of the microbiological as well as the sensorial quality of RTE pineapple through the application of spectroscopy and multispectral imaging technologies (sensors). The applied analytical technologies (FTIR, NIR, FLUO, VIS and MSI) were coupled with different machine learning algorithms (linear and non-linear). The suitability of the different algorithm-analytical technology combinations were comparatively evaluated. Additionally, the capabilities and the limitations provided by two commercially statistical data analytics tools were also explored.

In **Chapter 5,** spectroscopy-based sensors (FTIR, VIS, NIR and MSI) were used for the assessment of the microbiological quality of RTE rocket and baby spinach salads stored under passive and active MAP conditions. For model development, different machine learning algorithms (PLSR and SVR) and two distinct data partitioning approaches were performed. The applicability of the different sensors and machine learning approaches were compared for the purpose of choosing the optimum combination in terms of model performance.

Chapter 6 concerns the evaluation of the microbiological spoilage of fresh oyster mushrooms (*Pleurotus ostreatus*) using spectroscopy-based sensors. Two analytical technologies (FTIR and MSI) were applied and PLSR algorithm was utilized for model development.

Chapter 7 concludes the present thesis, providing a compilation and discussion of its main findings as well as future research perspectives.

SECTION A

Metagenetic characterization of microbial communities associated with fresh-cut produce spoilage

Chapter 2

Metagenetic characterization of microbial communities associated with ready-to-eat pineapple during storage under different temperature conditions

Part of this chapter is published in Food Microbiology 97 (2021) 103736 (Appendix II) 'Evolution of fungal community associated with ready-to-eat pineapple during storage under different temperature conditions'.

***Evanthia Manthou**, Gwendoline Coeuret, Stephane Chaillou, George-John E. Nychas*

Chapter 2

Abstract

The international market of fresh-cut produce commodities has witnessed dramatic growth in recent years, stimulated by consumer's demand for healthy, nutritious and convenient foods. One of the main challenging issues for the quality of these products is the potential microbiological spoilage that can significantly reduce their shelf-life. The complete identification of the microbiota of fresh-cut produce commodities together with the evaluation of environmental factors' impact on microbial composition is of primary importance. Therefore, the fungal and bacterial communities associated with the spoilage of ready-to-eat (RTE) pineapple were characterized using a metagenetic amplicon sequencing approach, based on the ITS2 and *gyrB* regions, respectively. The results revealed a significant variability in fungal species composition among different batches of RTE pineapple. The initial microbiota composition was the main influential factor and determined the progress of spoilage. Temperature and storage time were the secondary factors influencing spoilage, with their impact depending on the initial prevalent fungal species. Contrarily, both temperature and batch factors did not significantly influence the bacterial pineapple communities. The results strongly suggest that further large-scale sampling of RTE pineapple production should be conducted in order to assess the full biodiversity range of fungal community involved in the spoilage process and for unravelling the impact of important environmental factors on the initial microbiota.

2.1. Introduction

Fresh-cut market has grown dramatically in recent years, as a result of changes on consumers' attitude. Ready-to-eat (RTE) fruits and vegetables fulfil the growing demand for healthy, convenient and minimally processed food products (Gorni et al., 2015; Qadri et al., 2015). However, the quality and safety assurance of these new types of fresh products is a major challenge for the fresh-cut industry and requires full involvement and increasing investigations of food scientists (Padrón-Mederos et al., 2020).

Fresh-cut fruit and vegetable products have a limited shelf-life due to accelerated physiological and biochemical changes occurring during their processing and storage (Di Egidio et al., 2009; Torri et al., 2010; Zhang et al., 2014). Indeed, processing treatments render such products prone to both spoilage and pathogenic microorganisms (Leff & Fierer, 2013; Qadri et al., 2015). Various studies underline the presence of phytopathogens and human pathogens, but also microorganisms with antagonistic properties against these pathogens, which have a significant influence on human health and products' quality (Gorni et al., 2015). Therefore, a better insight into the microbial community and its potential interactions in food-associated matrices is required to provide safe and high-quality food (Juste et al., 2008; Cao et al., 2017).

So far culture-dependent methods have been the gold standards in food microbiology, since they have led to the description of a number of habitats. However, they are extremely biased in their ability to unravel the microbial communities of complex matrices associated with food or environmental samples (Juste et al., 2008; Ercolini et al., 2013; Zhou et al., 2015; Edet et al., 2017). On the other hand, the development of next generation sequencing (NGS) techniques has enabled researchers to study food microbial ecology from broader and deeper perspectives. Recently, metagenetic and metagenomic approaches have resulted in improved understanding of food microbiota by providing a species- and strain-level characterization (Abdelfattah et al., 2016; Abdelfattah et al., 2018; Poirier et al., 2018; Cauchie et al., 2020). Most NGS food microbiota studies have focused on fermented foods of different origins and, to a lower extent, on fresh meat and seafood products. On the other hand, NGS microbiota studies concerning fruits and vegetables are limited and mainly focus on epiphytic microbial communities (Lopez-Velasco et al., 2010; Rastogi et al., 2012; Jackson et al., 2013; Dees et

al., 2014; Abdelfattah et al., 2016; Söderqvist et al., 2017; Saminathan et al., 2018; Angeli et al., 2019; Tatsika et al., 2019; EFSA, 2020).

Pineapple (*Ananas comosus*) is one of the most popular tropical fruit worldwide and it is commonly found in the fresh-cut market. However, little has been reported about the associated microbial community and its response to various environmental factors (Montero-Calderon et al., 2008; Di Cagno et al., 2010; Dos Santos Souza et al., 2019). According to the limited studies based on the pineapple diversity, fungi have the leading role in fresh-cut pineapple's spoilage. The fungal species, even the prevalent ones, reported in pineapple differ among various studies (Tournas et al., 2006; Di Cagno et al., 2010; Chanprasartsuk et al., 2010; Leneveu-Jevrin et al., 2020). Interestingly, the existing literature is based on earlier generation molecular methods combined largely with culture-dependent and recently with culture-independent techniques.

In the present study, a metagenetic amplicon sequencing approach, based on the ITS2 region, was used with the objective to assess the fungal communities associated with the spoilage of RTE pineapple. Additionally, characterization of bacterial communities in pineapple was also performed using *gyrB* amplicon sequencing. To our knowledge, metagenetic analysis has never been applied to pineapple's microbiota. Therefore, this work sheds light on the variability of RTE pineapple's spoilage microbiota and on how it is changing during the product's shelf-life under the influence of storage temperature.

2.2. Materials and Methods

2.2.1. Sample preparation and storage conditions

Four batches of fresh-cut pineapple were supplied by a local manufacturer in Athens. The pineapple was packed in PVC trays containing 220 g of fruit. The trays were transported to the laboratory within 24 hours from their production and stored at three different constant temperatures, namely 4, 8 and 12 ^{o}C, and under dynamic temperature conditions (8 h at 4 ^{o}C, 8 h at 8 °C and 8 h at 12 ^{o}C) in high precision (±0.5) programmable incubators (MIR-153, Sanyo Electric Co., Osaka, Japan). The incubation temperature was recorded every 15 minutes using electronic temperature devices (COX TRACER®, Cox Technologies Inc., Belmont, NC, USA). The first sampling was conducted at the time of the product's arrival to the laboratory

and also at 38 , 72 , 134 and 230 h of storage at 4, 8 ^{o}C and at the dynamic temperature conditions, while with regard to storage at 12 °C, the final sampling point was 134 h.

2.2.2. Microbiological analysis and pH measurements

The samples (25 g of pineapple) were aseptically transferred into a sterile Stomacher bag (Seward Medical, London, UK), diluted with 225 ml of Ringer buffer solution (Lab M Limited, Lanchashire, UK) and homogenized for 60 sec in a stomacher device (Lab Blender 400, Seward Medical). The appropriate serial decimal dilutions were prepared and the following microbial determinations were performed: total mesophilic microbial populations (total viable count, TVC) by the spread method on tryptic glycose yeast agar (Plate Count Agar, Biolife, Milan, Italy), after incubation of plates at 25^{o}C for 72 h; yeast and moulds by the spread method on Rose Bengal Chloramphenicol agar (RBC, Lab M Limited) and incubation at 25^{o}C for 3-5 days. The results were expressed as the average (± standard deviation, *n*=4) log colony forming units per gram (log CFU/g) of the four pineapple batches.

The pH values of fruit samples were measured with a digital pH meter (RL150, Russell pH Cork, Ireland) with a glass electrode (Metrohm AG, Herisau, Switzerland).

2.2.3. DNA extraction of the microbiota recovered from plates

After the enumeration of the microbial populations, appropriate countable RBC plates were selected. All the colonies present on the surface of each plate were suspended in 2 ml aliquots of quarter-strength Ringer's solution (Lab M Limited), transferred with a sterile pipette in a 2-ml vial, and stored by freezing at -80^{o}C after addition of sterile glycerol (50% v/v).

Microbial DNA was extracted as previously described by Hoffman and Winston (1987) with slight modifications. Briefly, 0.3 ml of lysis solution and 0.3 ml of phenol/chloroform were added in the microbial pellets obtained after centrifugation (for the removal of glycerol). The solution was transferred in tubes with 0.3 g of glass beads which were then vortexed for 4 min. The tubes were centrifuged for 2 min at 13,000 rpm (at room temperature) and the supernatant was carefully transferred in 1.5-ml tubes. Absolute ethanol was added in 800 μl portions in each tube, and the tubes were centrifuged at the same conditions. After centrifugation, 1 ml aliquots of 70% ethanol were added in each tube, the tubes were

centrifuged once more, and the supernatants were discarded. The DNA was resuspended in 100 μl of TE buffer solution and stored overnight at 4 ºC. Before further analysis, 1 μl of RNase was added and the DNA was incubated for 15 min at 37 ºC.

2.2.4. DNA extraction of the pineapple microbiota

Ten grams of each pineapple sample were homogenized with 20 ml of Ringer's solution (Lab M Limited) in filter Stomacher bag (Interscience, St-Nom, France) in a stomacher device as described previously. Then, 20 ml of the homogenate were collected in 50-ml Falcon tubes (SARSTEDT AG & Co. KG, Germany) and centrifuged (Heraeus Multifuge 1S-R, Thermo Electron Co., US) at 8000 × g for 20 min at 4 ºC. The supernatant was discarded, the microbial pellet was washed with 20 ml of distilled-dionized water and centrifuged (Heraeus Fresco 21, Thermo Scientific, US) once more at the same conditions. The cells were resuspended in 1.7 ml of sterile ultrapure water, transferred in 2-ml eppendorfs (SARSTEDT AG & Co. KG), and centrifuged at 17000 x g for 10 min at 4 ºC. The supernatant was discarded and the microbial cells were stored at -80ºC. Fungal DNA was extracted according to the protocol described above for plates, while bacterial DNA was extracted with the DNeasy PowerSoil Kit (Qiagen, Hilden, USA) according to manufacturer's instructions.

2.2.5. Barcoding PCRs and Illumina Miseq PCR

2.2.5.1. ITS2

Amplicon libraries were constructed following two rounds of PCR amplification. The first amplification of the ITS2 rRNA gene was performed with the primers ITS3 (5'-GCATCGATGAAGAACGCAGC-3') and ITS4 (5'-TCCTCCGCTTWTTGWTWTGC- 3'). The final primer concentration used was 10 μM. Forward and reverse primers carried the Illumina 5'-CTTTCCCTACACGACGCTCTTCCGATCT-3' and the 5'-GGAGTTCAGACG TGTGCTCTTCCGATCT-3' tails, respectively. The first round of PCRs was performed with the high-fidelity AccuPrime Taq DNA polymerase system (Invitrogen, Carlsbad, USA) and 5 μL of microbial DNA. The cycling conditions were: 94ºC for 1 min, followed by 30 cycles of amplification at 94ºC (60 sec), 55ºC (60 sec), and 72ºC (60 sec), with a final extension step of 10 min at 72ºC. The amplicon size, quality, and quantity of the amplified DNA were checked on a DNA1000 chip (Agilent Technologies, Paris, France). Then, the second Miseq PCR and

the Miseq sequencing was conducted with V3 illumina MiSeq kit as described in Poirier et al. (2018). Raw read sequences were deposited at the Sequence Read Archive under the Bioproject number PRJNA665125 and the accession numbers SAMN16242305 to SAMN16242366.

2.2.5.2. gyrB

Two rounds of PCR amplification took place for the construction of amplicon libraries. The first amplification of the 250 bp region of *gyrB* was performed with the primers F64 (5'-MGNCCNGSN ATGTAYATHGG -3') and R353 (5'-ACNCCRTGNARDCCDYCNGA- 3'). Forward and reverse primers carried the Illumina 5'-CTTTCCCTACACGACGCTCTTCCG ATCT-3' and the 5'-GGAGTTCAGACGTGTGCTCTTCCGATCT-3' tails, respectively. The first round of PCR was performed again with the high-fidelity AccuPrime *Taq* DNA polymerase system (Invitrogen, Carlsbad, USA), 20 μM final primer concentration and 5 μL of microbial DNA. Amplification of *gyrB* was performed with the following cycling conditions: 94 ^{o}C for 2 min, followed by 35 cycles of amplification at 94 ^{o}C (30 sec), 55 ^{o}C (60 sec), and 68 ^{o}C (90 sec), with a final extension step of 10 min at 68 ^{o}C. The amplicon size, quality, and quantity of the amplified DNA were checked on a DNA1000 chip (Agilent Technologies, Paris, France). For each pineapple batch, the amplified DNA samples derived from the different storage times (38, 72, 134, 230 h) were pooled together in terms of temperature and were further analysed together with the initial bacterial communities (0 hour). The samples derived from dynamic temperature conditions were not sequenced. The second Miseq PCR and the Miseq sequencing were conducted as described in Poirier et al. (2018).

2.2.6. Quality filtering of reads and taxonomic assignment of Operational Taxonomic Units (OTU)

Raw sequencing reads were imported into the FROGS (Find Rapidly OTUs with Galaxy Solution) pipeline (Escudie et al., 2017) for quality control and assembly into OTU. Roughly, the pipeline was as follows: quality-filtered ITS2 and *gyrB* paired-end sequences were merged into contigs with VSEARCH v2.15.0 (Rognes et al., 2016) using 0.1 mismatch rate in the overlapped region. Only amplicons with size above 150 bp but no longer than 500 bp and 350 bp were kept for fungi and bacteria, respectively. Merged amplicon sequences were

dereplicated and clustered using SWARM v3.0.0 algorithm (Mahe et al., 2015) with a distance threshold of 3. Chimeras were removed with VSEARCH v2.15.0. The resulting sequences were filtered for spurious OTUs by keeping only those with at least 0.01% of relative abundance within the whole dataset (Auer et al., 2017). Taxonomic assignment of OTUs was performed using the UNITE 6.1 ITS2 for fungi, the *gyrB_03_2019* for bacteria as reference databases (Nilsson et al., 2019 https://unite.ut.ee/), as well as the Blastn+ algorithm (Camacho et al., 2009).

2.2.7. Analysis of alpha and beta diversity

Fungal diversity was analysed using the R package Phyloseq (McMurdie & Holmes, 2013). OTU abundance was normalized using the median sequencing depth of all samples. Analyses of alpha and beta diversity were carried out using standard or custom Phyloseq command lines.

2.3. Results

2.3.1. Growth of dominant fungal microbiota is temperature dependent

A comparative analysis of TVC and fungal counts on four independent pineapple batches revealed that fungi and mostly yeasts are the main component of the cultivable microbiota over storage (Figure 2.1).

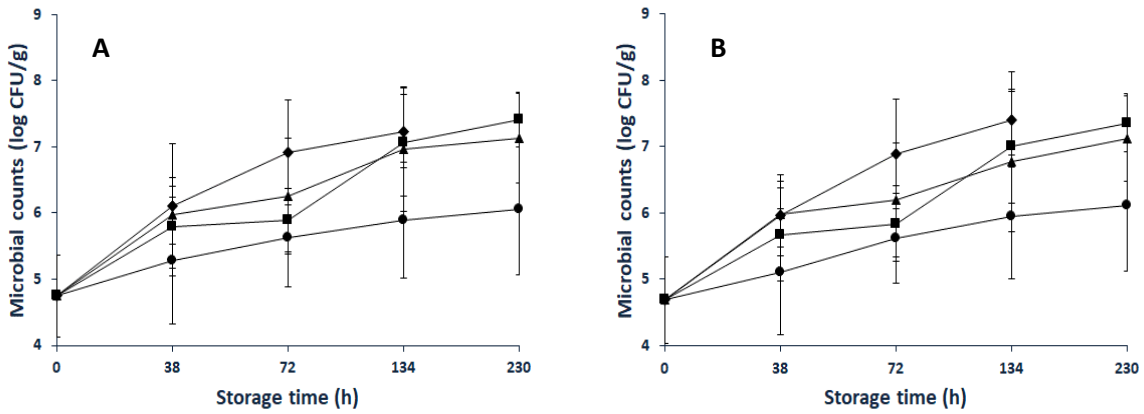


Figure 2.1. Total viable counts (TVC) (A) and fungal populations (B) in ready-to-eat pineapple during storage at 4 ºC (●), 8 ºC (■), 12 ºC (◆), and dynamic temperature conditions (▲). The microbial populations are expressed as means ± standard deviations (n=4).

The yeast populations almost coincide with those of total mesophiles throughout storage and for all the studied temperatures. As expected, the growth of the fungal population was faster at the highest temperatures. The initial level of fungi (mean ± standard deviation, n=4) was 4.69 ± 0.65 log CFU/g, and reached a final average level of 7.36 ± 0.44 and 7.41 ± 0.72 log CFU/g at 8 and 12 ºC, respectively. Storage at 4 °C revealed more stringent than the three other temperature conditions on growth. In this case, the fungal population reached only 6.11 ± 0.99 log CFU/g after 230 h. The microbial growth monitored during dynamic temperature conditions resembled that recorded at 8 ºC. As far as the pH is concerned, no considerable differences in pH measurements were found among the different temperatures and during storage (Table 2.1).

Table 2.1. The initial and final pH values of ready-to-eat pineapple during storage at 4, 8, 12 ºC and dynamic temperature conditions.

	pH	
Storage temperature	**Initial value**	**Final values**
4 ºC	3.32 ± 0.25	3.60 ± 0.07
8 ºC		3.56 ± 0.06
12 ºC		3.66 ± 0.06
Dynamic		3.64 ± 0.11

2.3.2. Fungal OTU richness and alpha-diversity is different among pineapple samples and cultivation media.

Pineapple and plates samples were sequenced at an average of 45.067 ±22.083 reads. The rarefraction curves (Figure 2.2) performed on quality-filtered reads indicated that sequencing depth was sufficient for all the samples tested in order to assess the OTU richness

present in pineapple and plates. Subsequently, it was investigated whether cultivation methods were underestimating the level of fungal population during storage. The fungal diversity by ITS2 amplicon sequencing between DNA extracted directly from pineapple samples and from the fungal population that grew on agar plates was compared. As shown in Table 2.2 and also indicated by Figure 2.3, the fungal OTU richness was significantly ($p<0.01$) lower on plates.

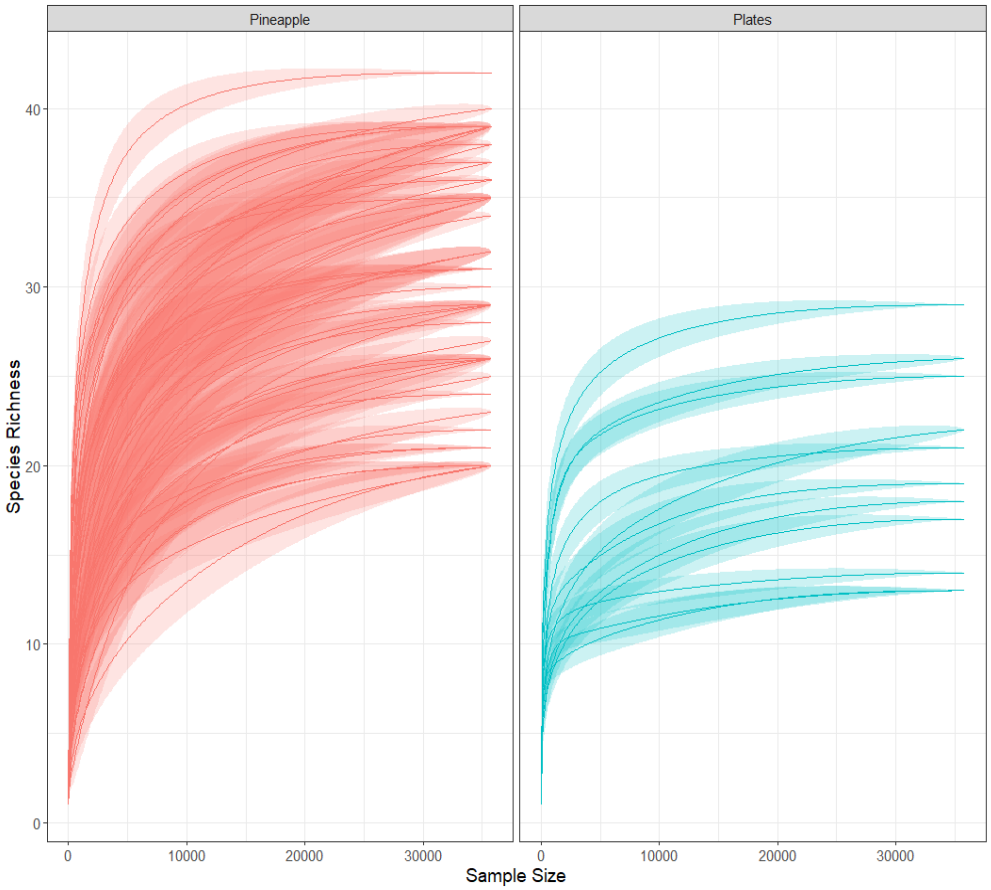


Figure 2.2. Rarefaction curves obtained from ITS2 amplicon sequencing of the pineapple and plate samples communities. The x-axis represents the sequencing depth in number of reads, while the y-axis represents an estimation of the OTU richness detected.

Moreover, Figure 2.3 shows how the 33 fungal species detected by non-cultural metagenetic analysis could be detected and quantified from the microbiota recovered from agar plates. In general, the detection of most species from the Basidiomycota phylum was unsuccessful in comparison to species from the Ascomycota phylum. In addition, detection of *Fusarium* species was also strongly biased on plates, in particular *Fusarium circinatum* was

highly abundant in pineapple samples compared to plates. Therefore, to avoid any bias in our analysis, only data from pineapple samples were used further.

Table 2.2. Fungal OTU richness (merged at different taxonomic levels) among pineapple samples and agar plate samples.

	Number of genera	Number of species	Number of OTUs
Pineapple samples	22	33	47
Plate samples	18	25	39

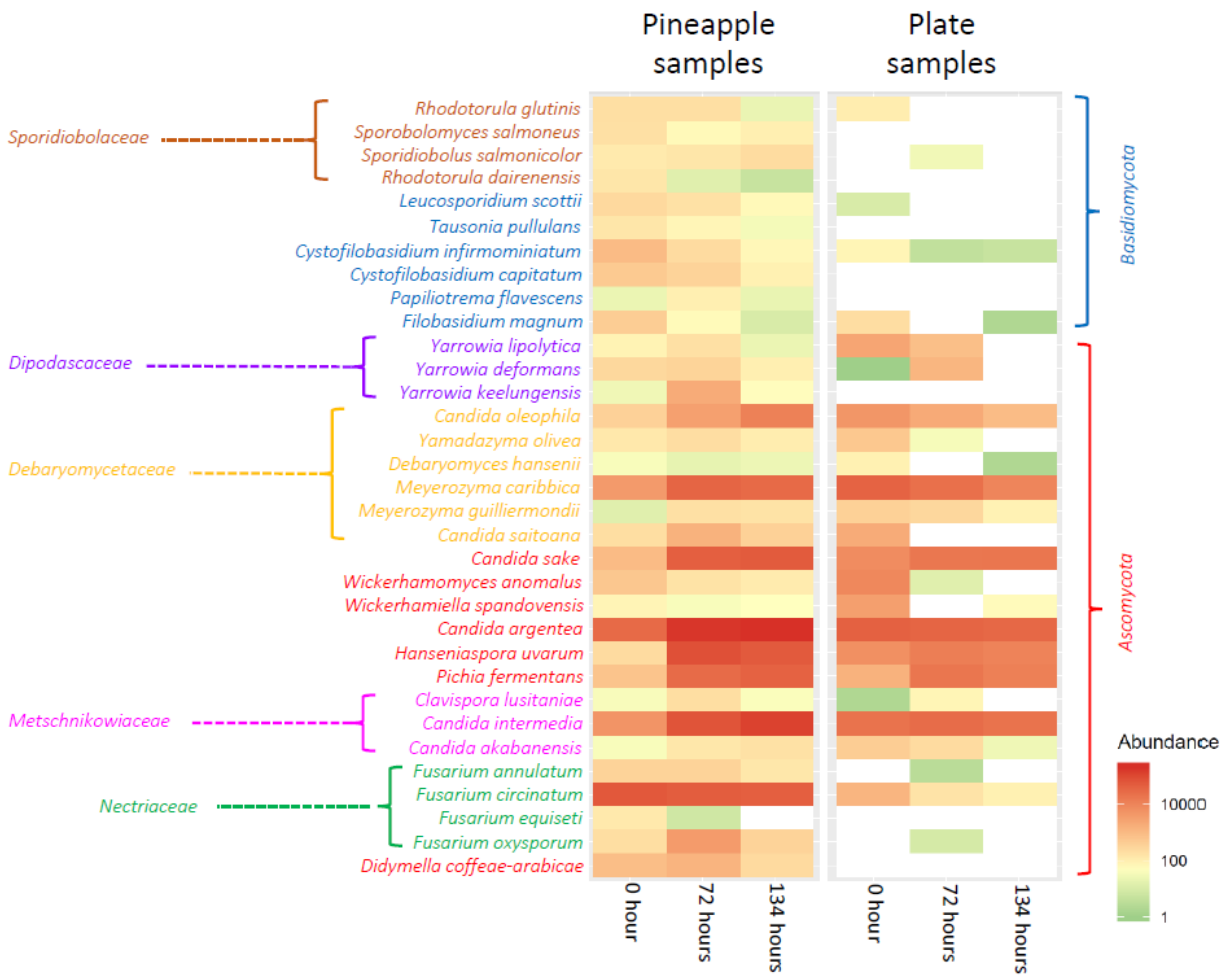


Figure 2.3. Heatmap showing the comparison among fungal species' relative abundances detected by non-cultural methods (pineapple samples) and by cultural methods (plate samples).

Each column shows the average relative abundance of the 33 species after merging of the various samples from the four pineapple batches and for the different storage temperatures. Main fungal families and phyla are depicted by bracklets.

2.3.3. Effect of temperature and time of storage on fungal diversity

The effect of different temperatures and storage times was analysed on fungal richness after merging OTUs at species-level (Figure 2.4). The species' richness was significantly ($p<0.01$) higher in pineapple samples stored at 4 ºC. The samples stored at dynamic conditions and 8 ºC followed, while the samples stored at the highest temperature (12 ºC) had the lowest number of species. The fungal richness was also comparable for the different storage times ($p<0.01$). At the beginning of the storage (time-zero), the diversity was higher compared to all the other storage times. Although the species richness decreased over time, the species' number at 230 h did not follow the same declining course. This observation is not unexpected, since the samples at 230 h came exclusively from storage at 4 ºC. The corresponding samples (at 230 h) stored at 8 ºC and dynamic conditions were not successfully sequenced.

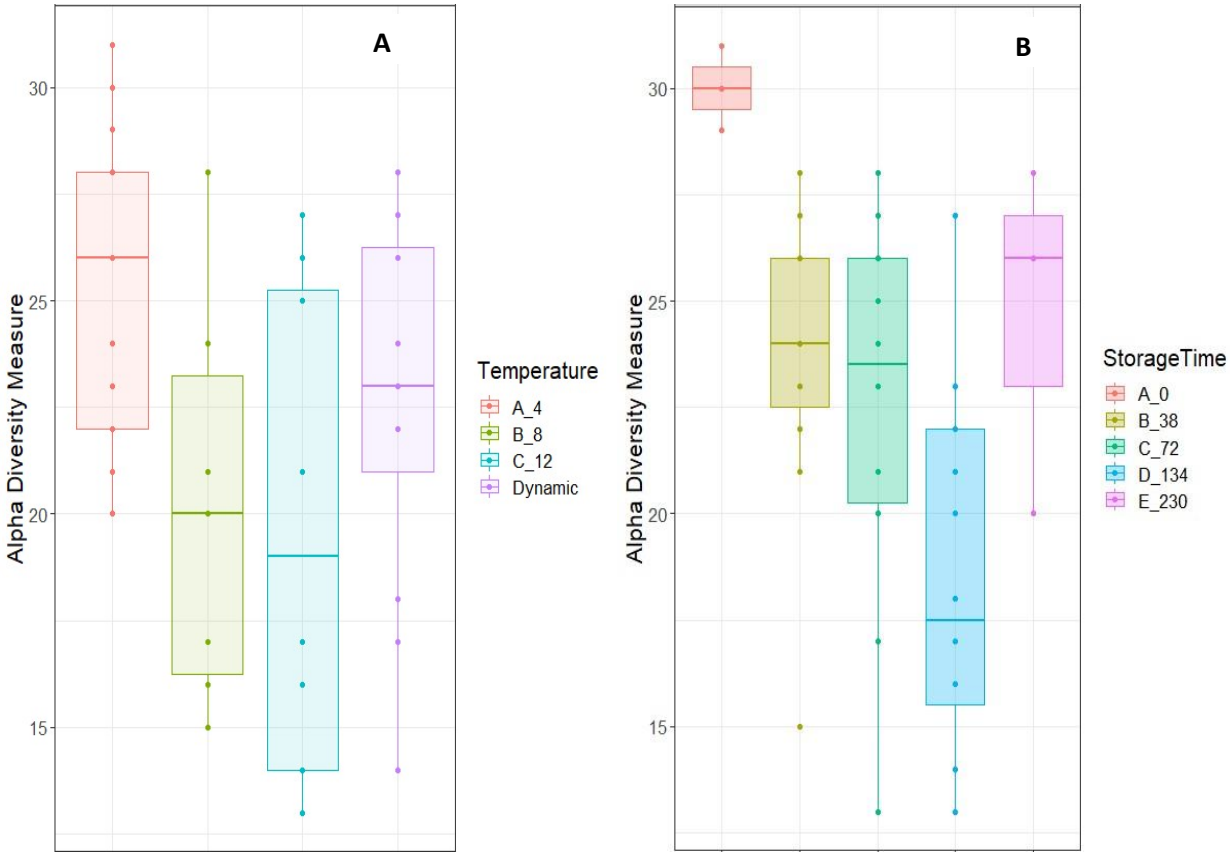


Figure 2.4. Fungal richness in ready-to-eat pineapple samples. The box plot shows the number of species in samples of different temperatures (A) and time (B) of storage. The boxes represent the interquartile range between the first and third quartiles and the vertical line inside the boxes is the median obtained from the samples analysed per condition. **Note**: The storage time of 230 h corresponds exclusively to samples stored at 4 ºC, since the corresponding samples (at 230 h) stored at 8 ºC and dynamic conditions were not successfully sequenced.

Non-Metric Multidimensional Scaling (NMDS) on Bray-Curtis distances were also performed to statistically compare the fungal diversity within the samples of different temperature and time of storage. In all cases, communities recovered from a given temperature or storage time did not cluster together on the factorial plane (data not shown). On the other hand, there was a discrete clustering among samples of the different batches of pineapple, presented in Figure 2.5.

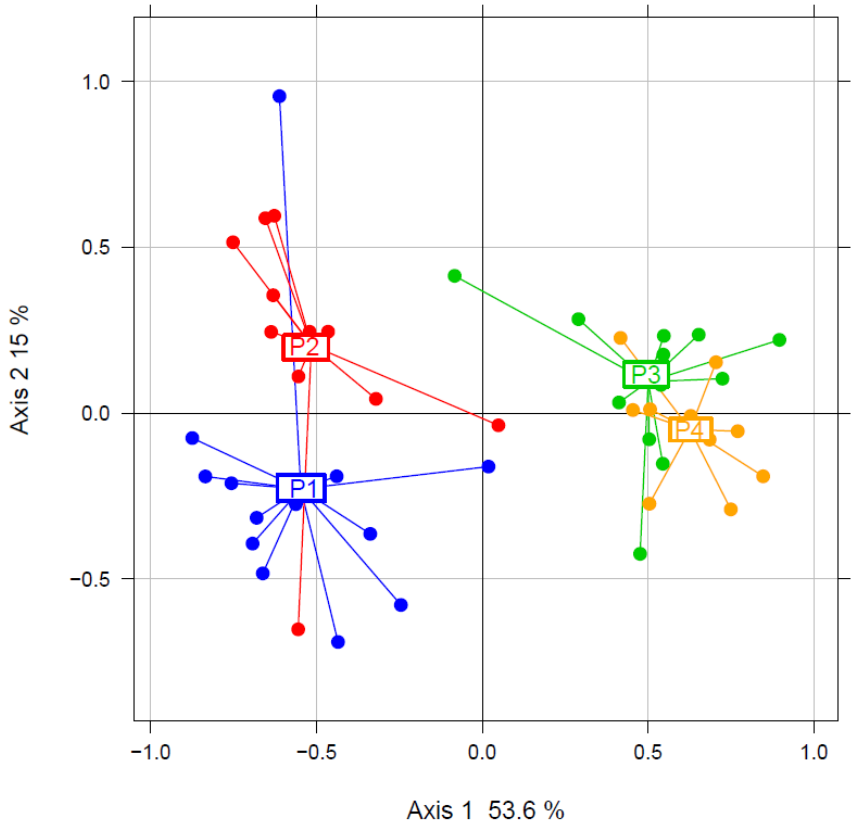


Figure 2.5. Non-Metric Multidimensional Scaling (NMDS) based on Bray-Curtis distances among fungal communities of the four ready-to-eat pineapple batches. The points represent

the fungal communities of the various samples, while the lines represent the rank orders (distances) among them. P1, P2, P3 and P4 correspond to batch 1, batch 2, batch 3 and batch 4.

The fungal diversity of most of the samples from batch P1 and P2 was distinct of that of batches P3 and P4 together. The differences on fungal diversity among the four batches are presented in Figure 2.6, where the composition plot of relative abundances is illustrated according to Bray-Curtis hierarchical clustering of pineapple samples. A phylogenetic tree of the different fungal species based on the ITS2 sequences is illustrated in Figure 2.7. Pineapple samples from batches P3 and P4 had a quite similar fungal community dominated by two phylogenetically related species, namely *Candida argentea* and *C. sake* from the family *Saccharomycetales_incertae_sedis*. Most samples from batch P4 can be distinguished from those of batch P3 with the presence of *Hanseniaspora uvarum,* which is also closely related phylogenetically to the aforementioned *Candida*. Pineapple samples from batches P1 and P2 displayed a set of completely different fungal communities dominated by species from the phylogenetically related *Nectriaceae* (*Fusarium*) and *Metschnikowiaceae* (*Clavispora/ Candida*) families. *Candida intermedia* and *Fusarium circinatum* were the most abundant species in samples of batch P2, whereas some samples from batch P1 showed higher level of diversity with *Pichia fermentans* (*Pichiaceae*) and *Meyerozyma caribbica* (*Debaryomecytaceae*).

At batch level, the effect of storage time and temperature varied across the different fungal species. Interestingly, the initial composition of pineapple microbiota had a great impact on the evolution of spoilage at different temperatures and storage time. Therefore, situations varied from one batch to another. Starting with batch P1 (Figure 2.8), it was observed that both temperature and storage time drove a strong change on the fungal community composition. Temperature was the most influential parameter separating samples stored at 4 °C from those stored at 12 °C with, in both cases, a gradual change over longer storage time. Samples stored at the intermediate temperature of 8 °C, clustered at intermediate positions between samples stored at 4 °C and 12 °C. Only samples stored at dynamic temperature conditions scattered randomly with no logical order. Among the most striking changes, it was observed that *F. circinatum* had high abundance (>75%) at time-zero of storage, while following the next stages

of storage *C. intermedia* or *P. fermentans* finally dominated depending on the temperature. Specifically, *C. intermedia* succeeded to dominate in samples stored at 12 and 8 ºC from the early stages of storage, while *P. fermentans* in samples stored at 4 ºC at the final stages (134 and 230 h). On the other hand, *M. carribica* was able to dominate only at the middle of storage (72 h) for 8 °C and dynamic conditions. Concerning batch P2 (Figure 2.9), a trend similar to that observed in batch P1 was observed with, at time-zero, a large dominance of *F. circinatum* which was progressively replaced by *C. intermedia* at high storage temperatures and by *C. argentea* at 4 °C. As shown in Figures 2.10 and 2.11, the fungal communities from samples of batches P3 and P4 were not affected significantly by temperature and storage time. Unlike pineapple samples from batches P1 and P2, the fungal community of batch P3 was covered by the great dominance of *C. argentea* throughout storage at all studied temperatures. *C. sake*, which was the second most abundant species for the most of the samples, was not specifically affected by temperature or time in P3. However, in the case of samples from batch P4, a slight impact of temperature and storage time on *H. uvarum* and *C. argentea* was noticed. Although there is no information for the initial composition (due to unsuccessful sequencing of 0-h sample), *H. uvarum* prevailed at 8 and 12 °C in the middle of storage (72 and 134 h), while the latter prevailed at 4 °C for the two storage times and at 12 °C and dynamic conditions at the early stages of storage (38 h). *C. sake* became more abundant during storage at dynamic conditions and finally prevailed at 134 h.

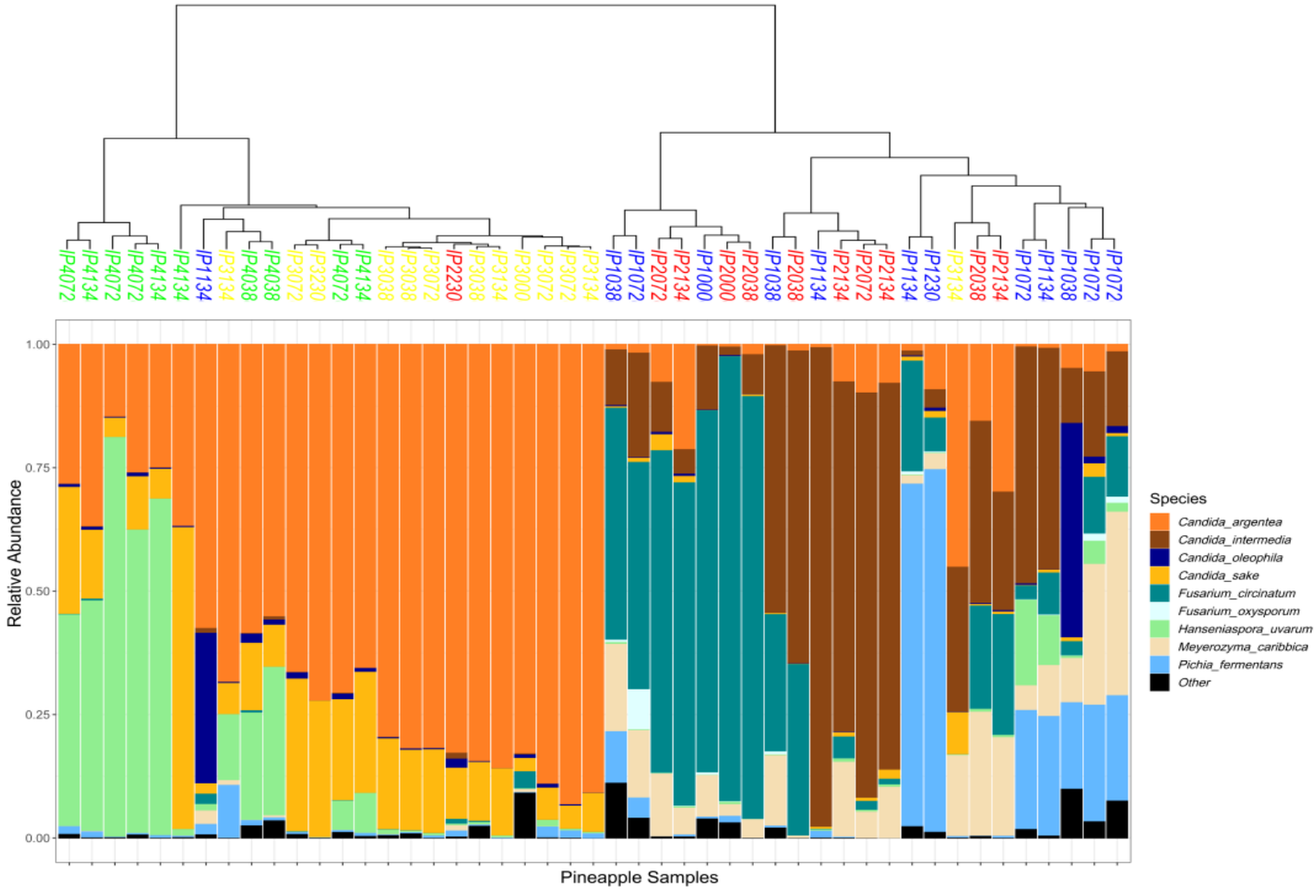


Figure 2.6. Composition plot showing the relative abundances of the nine main Ascomycota species found in ready-to-eat pineapple samples. On the top: hierarchical clustering of batches samples according to Bray-Curtis distance and ward algorithm (blue for P1, red for P2, yellow for P3 and green for P4 as shown in Figure 2.5.

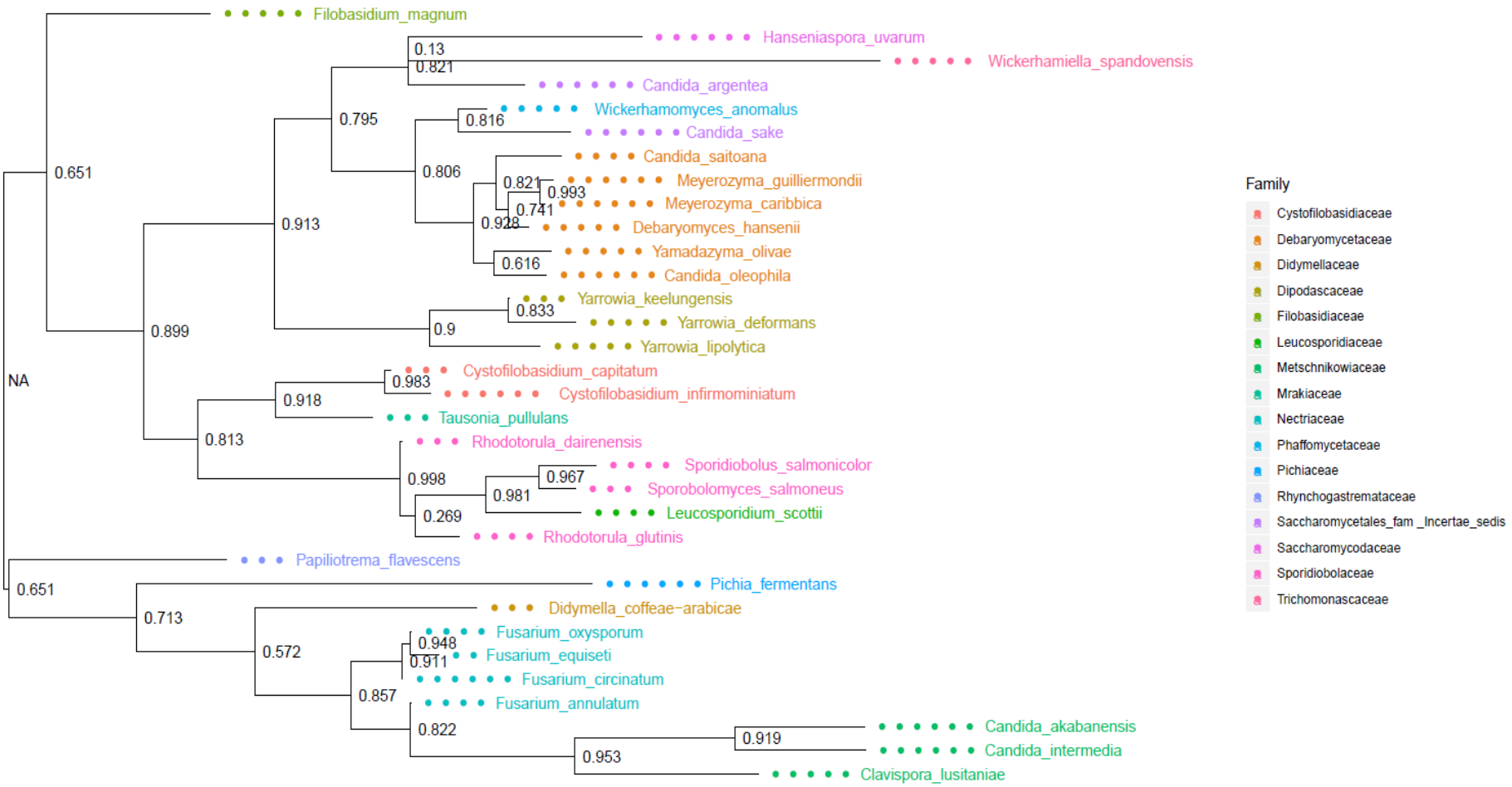


Figure 2.7. Neighbour-Joining phylogenetic tree of the different species based on the ITS2 sequences of ready-to-eat pineapple samples. Bootstrap values are indicated on the main nodes.

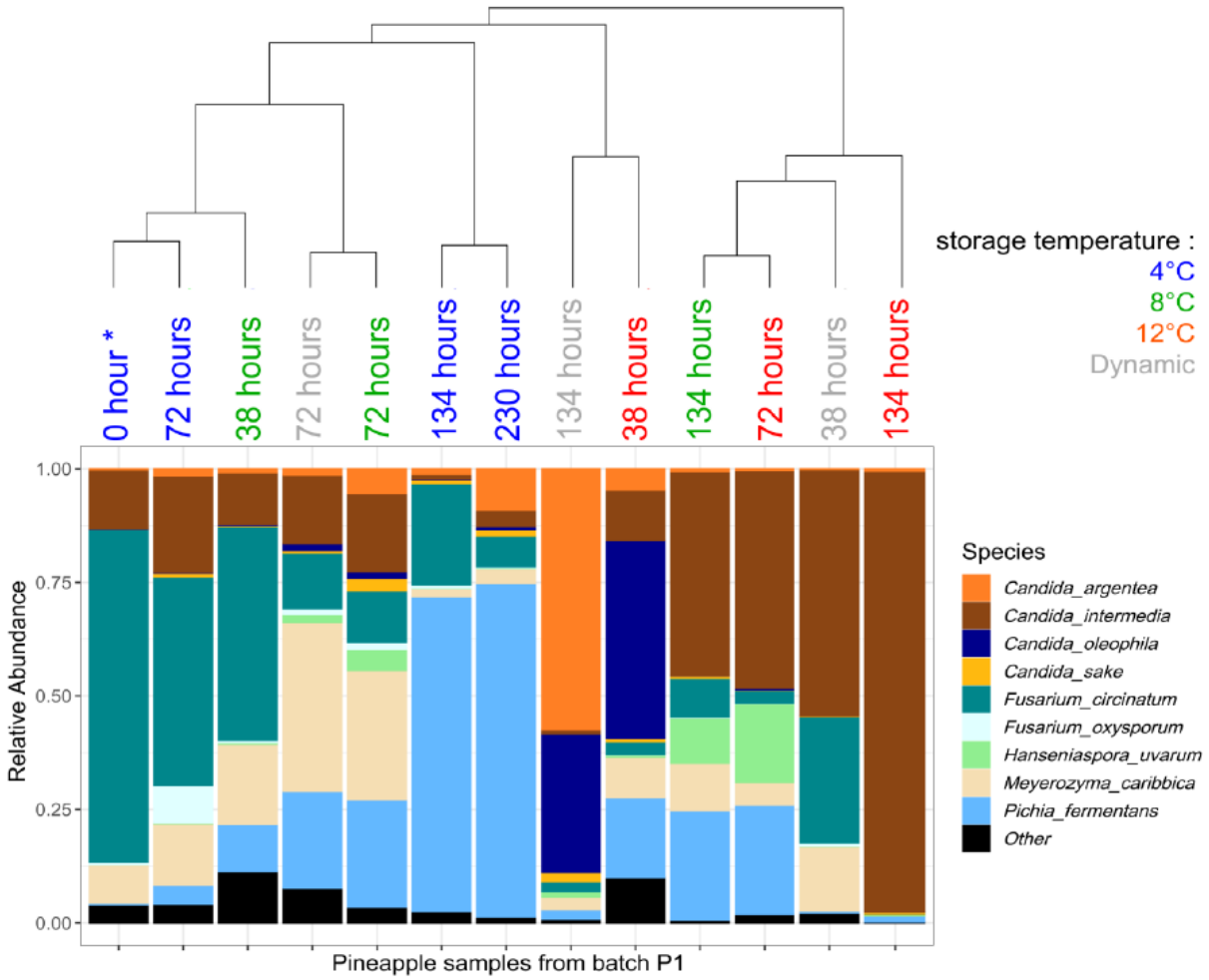


Figure 2.8. Impact of storage time and temperature on fungal species composition of ready-to-eat pineapple samples from batch P1. Samples are ordered from left to right according to Bray-Curtis distance. The asterisk (0 h) indicates the initial analysis before packaging and storage at any other conditions.

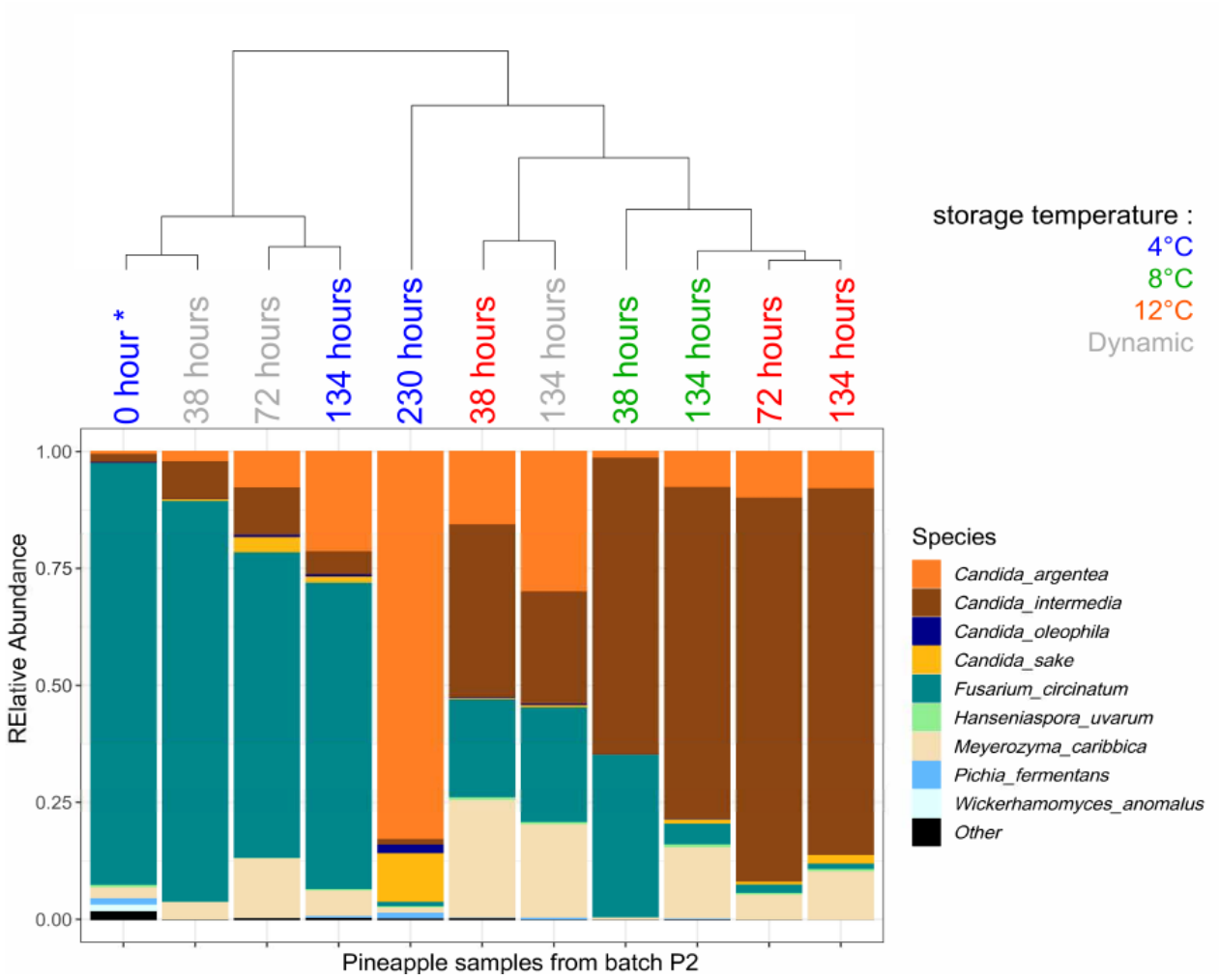


Figure 2.9. Impact of storage time and temperature on fungal species composition of ready-to-eat pineapple samples from batch P2. Samples are ordered from left to right according to Bray-Curtis distance. The asterisk (0 h) indicates the initial analysis before packaging and storage at any other conditions.

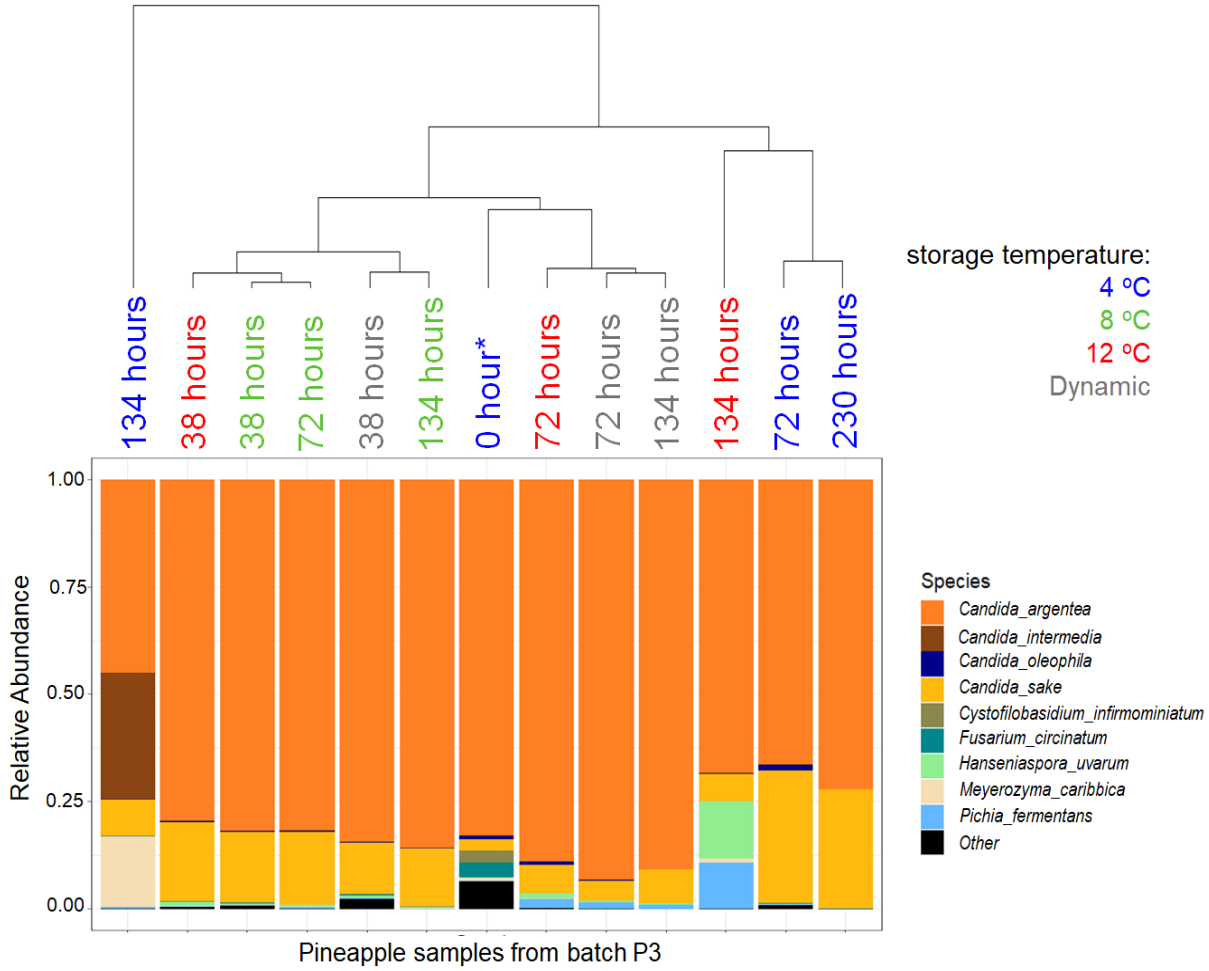


Figure 2.10. Impact of storage time and temperature on fungal species composition of ready-to-eat pineapple samples from batch P3. Samples are ordered from left to right according to Bray-Curtis distance. The asterisk (0 h) indicates the initial analysis before packaging and storage at any other conditions.

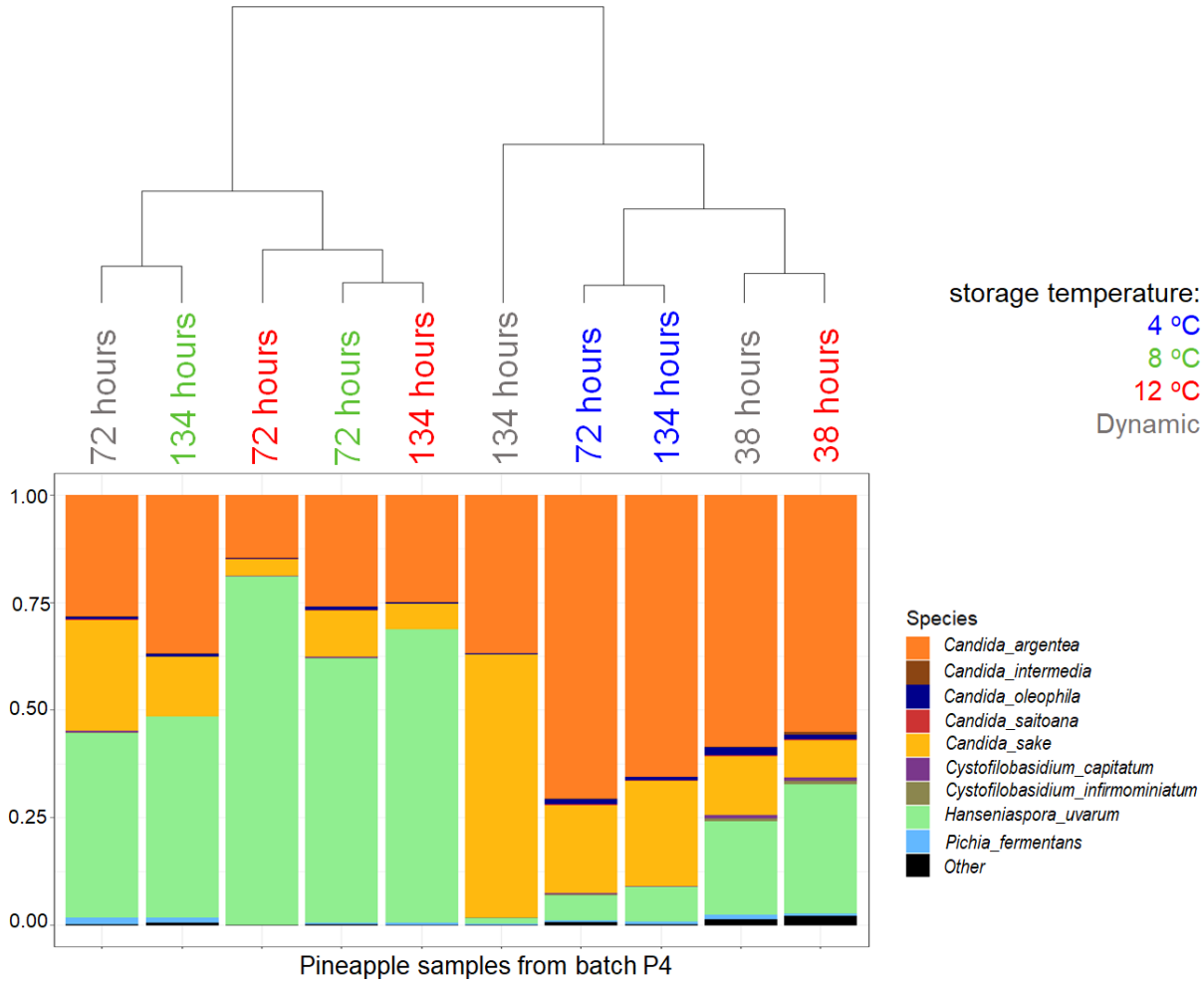


Figure 2.11. Impact of storage time and temperature on fungal species composition of ready-to-eat pineapple samples from batch P4. Samples are ordered from left to right according to Bray-Curtis distance.

2.3.4. Bacterial diversity and the effect of temperature

As far as the bacteria are concerned, results derived from preliminary microbiological analysis indicated that *Pseudomonas* spp. and bacteria of the Enterobacteriaceae family were often below the detection limit or not more than 2 log CFU/g throughout storage at all the studied temperatures (data not shown). However, *Pseudomonas* colonies grown on agar plates were not macroscopically the typical ones (not shiny but dehydrated colonies), even though the results from the oxidase, catalase and Gram stain tests for some of them that were randomly selected and tested, were often the expected ones. Moreover, the population of lactic acid bacteria (LAB) was not easily detectable with common microbiological analyses due to the dominance of yeasts in the MRS medium (even when cycloheximide was added to

the medium). With regard to the above difficulties and the bias of conventional microbiological methods, this observation led to the investigation of the bacterial pineapple microbiota.

The rarefraction curves (Figure 2.12) performed on quality-filtered *gyrB* reads indicated that sequencing depth was sufficient for all samples tested. In total, 158 OTUs were detected at the species taxonomic resolution. As it was previously mentioned, the bacterial samples of each pineapple batch have been pooled according to temperature, since it was initially hypothesized that temperature could possibly have a significant impact on microbial communities and also the subdominant bacterial populations would not considerably change during storage. The 0-h samples represent the initial community composition for each batch. Not surprisingly, the species richness and diversity of pooled samples was significantly ($p<0.05$) higher than samples derived from 0 h (data not shown).

The effect of the different temperatures on bacterial diversity was investigated after merging OTUs at species-level. The species' diversity was not significantly influenced by the temperature in terms of both richness and evenness. Indeed, the communities recovered from a given temperature did not cluster together using NMDS on Bray-Curtis and Unifrac distances. Similarly, there were no significant differences among samples of the different batches of pineapple (data not shown).

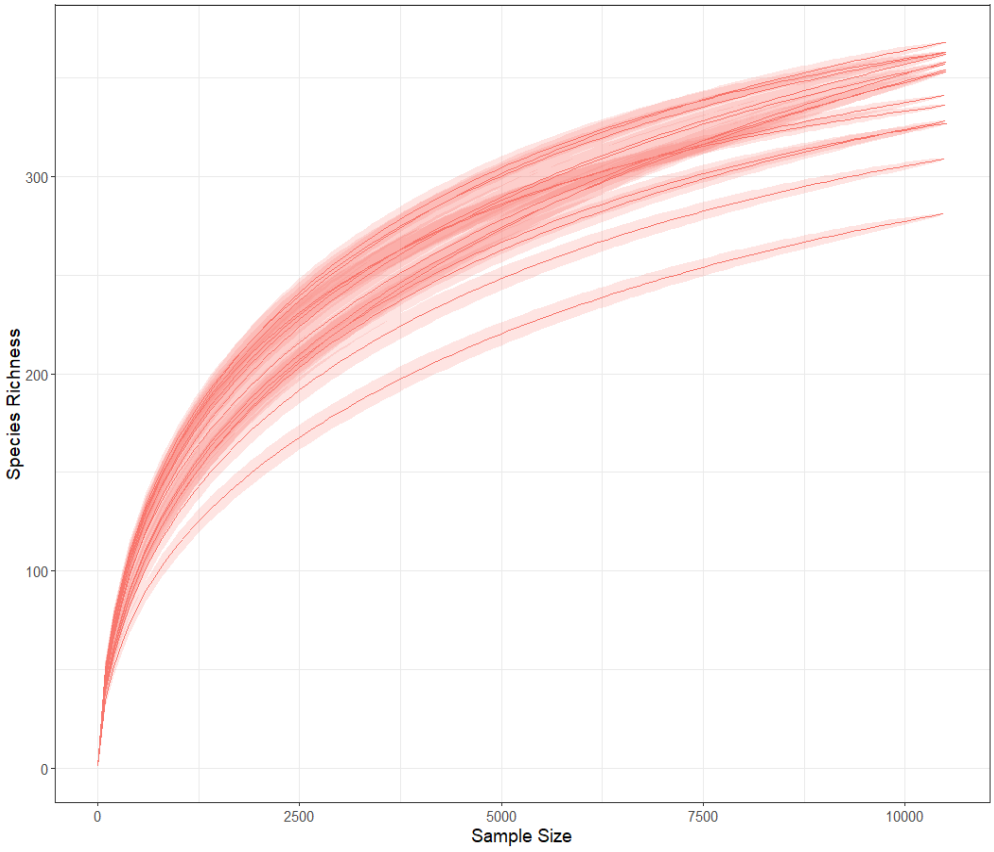


Figure 2.12. Rarefaction curves obtained from *gyrB* amplicon sequencing of the ready-to-eat pineapple communities. The x-axis represents the sequencing depth in number of reads while the y-axis represents an estimation of the OTU richness detected.

According to the composition plot (Figure 5.13), *Pseudomonas* was the most prevalent genus in pineapple bacterial microbiota. Among the *Pseudomonas* species, *Pseudomonas viridiflava* was the most abundant for all the samples. In some cases, this species showed the highest relative abundance among the 19 bacterial species illustrated in Figure 2.13. Members of the Enterobacteriaceae family as well as LAB also were present. Few samples derived from different batches and temperatures showed higher abundances of *Lecleria adecarboxylata*, *Lelliottia amnigena*, *Citrobacter freundii* and *Pectobacterium carotovorum* (belonging to the Enterobacteriaceae family) compared to the others. The same samples had also higher abundance of the lactic acid bacterium, *Lactococcus piscium*.

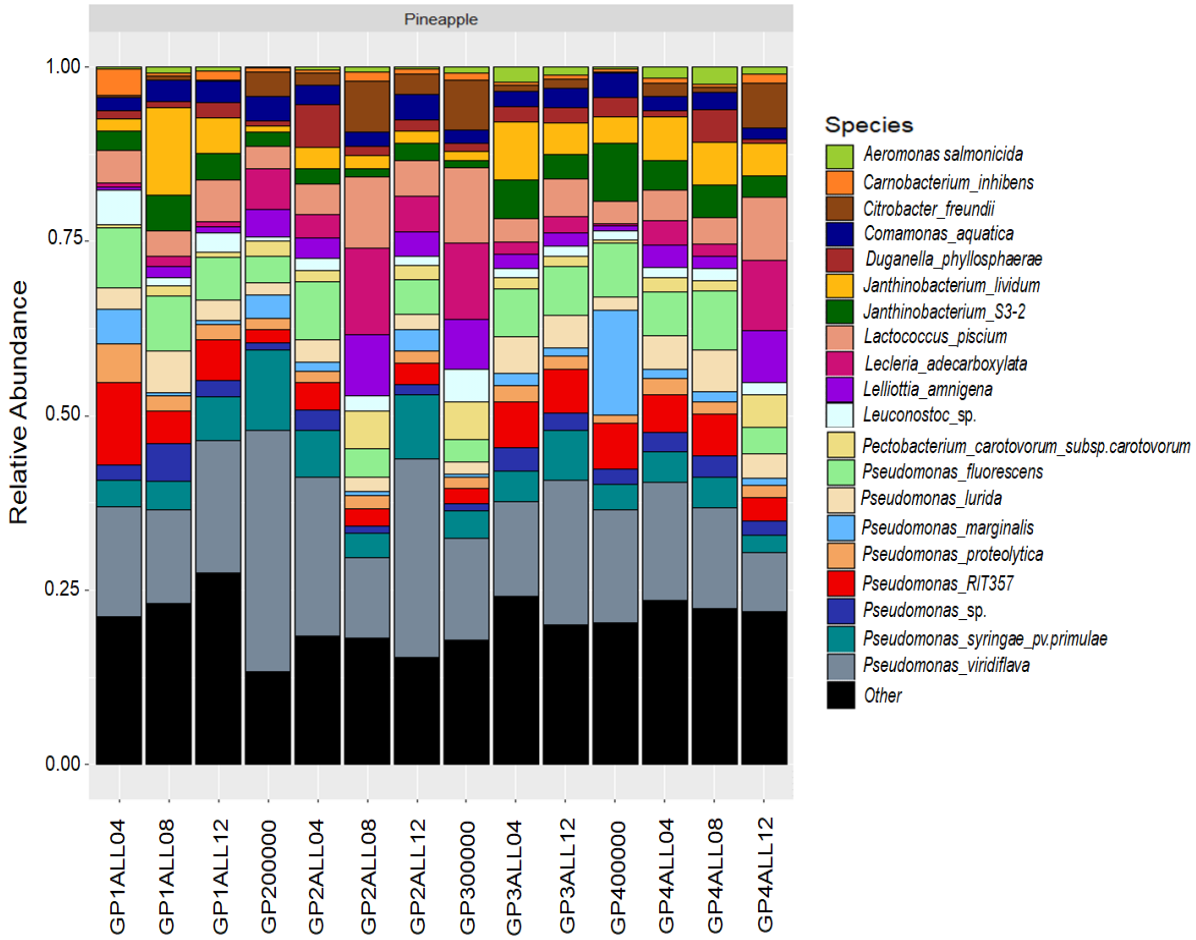


Figure 2.13 Composition plot showing the relative abundances of the 19 main bacterial species found in ready-to-eat pineapple samples (G: *gyrB*, P: pineapple, ALL: all storage times pooled, 04:4 ºC, 08:8 ºC, 12:12 ºC)

2.4. Discussion

The aim of the present study was to characterize the microbial community involved in the spoilage of RTE pineapple and determine the changes of the diversity when stored under different temperatures, using a metagenetic approach. The results demonstrated that fungi and mainly yeasts were the predominant spoilage microorganisms found in RTE pineapple. Not surprisingly, the high yeast population is likely attributed to the high level of sugars and the low pH of pineapple, which is ideal for their growth (Maciel et al., 2013; Da Cruz Almeida et al., 2018; Leneveu-Jenvrin et al., 2020). Similar levels of mesophilic populations and fungi in pineapple were also recorded in previous studies (Tournas et al., 2006; Montero-Calderon et al., 2008; Di Egidio et al., 2009; Leneveu-Jevrin et al., 2020). Two of these studies also pointed out significant differences in the initial and final microbial concentrations among distinct

product batches (Tournas et al., 2006; Leneveu-Jevrin et al., 2020). However, there is a lack of information about the populations of specific bacterial groups in fresh-cut pineapple, since the majority of the studies enumerate only the total microbial populations and the dominant yeasts. The results from the preliminary research revealed low populations of bacteria, but also various difficulties in determining specific bacterial groups. Contrarily, Di Cagno and colleagues (2010) reported initial LAB populations close to 6 log units. Moreover, Leneveu-Jevrin et al. (2020) reported initial Enterobacteriaceae populations ranging from 3.0 to 5.3 log CFU/g, while after storage at 4 ºC for 10 days, these populations reached maximum values of 7.8 log CFU/g.

Although cultural methods have been extensively used in food microbiology, they are considered extremely biased in their ability to capture the microbial diversity of complex environments (Ercolini et al., 2013; Cao et al., 2017; Edet et al., 2017). Consequently, a comparative analysis of the culture-dependent and culture-independent characterization of fungal community of pineapple was first attempted in the context of the present study. It was demonstrated that a whole phylum (i.e. Basidiomycota) was hardly detected in plates, while one of the most abundant species detected in one batch (*F. circinatum*) was unsuccessfully represented in the microbiota recovered from plates. This is indicating that non-cultural analytical methods should be used for further microbiota analysis of pineapple.

The effect of temperature and storage time on fungal diversity of pineapple samples was further analysed. Both factors had a significant impact; specifically, species richness decreased in the course of storage and with increasing storage temperature. However, when the fungal diversity within the samples was comparatively assessed for different temperatures and storage times, based on Bray-Curtis distances, no radical clear clustering could be evidenced under the conditions of this study. On the other hand, the analysis revealed a significant batch effect on fungal diversity and composition. Leneveu-Jenvrin et al. (2020) also observed that the pineapple communities clustered mostly according to batches and not to storage time. The high biological variability is a common observation in plant-origin products due to the strong impact of various factors such as cultivar, geographical region and agricultural practices on their microbiological quality status (Leff & Fierer, 2013).

In the present work, the two batches (P3 and P4) had similar fungal composition, since the phylogenetically related *C. argentea* and *H. uvarum* (only in batch P4) or *C. sake* were the

dominant species. On the other hand, most samples from P1 and P2 were dominated by the species *F. circinatum* or *C. intermedia* which are also phylogenetically related to each other. Moreover, some samples of P1 showed higher level of diversity, since *P. fermentans* and *M. caribbica* were present in high abundances. There are few previous studies on the fungal composition of pineapple, albeit not using NGS and being mainly based on culture-dependent techniques. Tournas et al. (2006) detected *Schwanniomyces polymorphus* (formerly known as *Debaryomyces polymorphus)*, *Candida pulcherrima*, *Pichia* spp. and in low abundances *Penicillium* spp., while Di Cagno et al. (2010) found *Meyerozyma guilliermondii* (formerly known as *Pichia guilliermondii)* as the only species recovered on plates. Chanprasartsuk et al. (2010) identified *M. guilliermondii* and *H. uvarum* as the main yeasts of fresh pineapple juices from different locations and countries. These two species were characterized both by culture-dependent and -independent techniques, using in the latter case denaturing gradient gel electrophoresis (DGGE). Zhang et al. (2014) used *Candida argentea, C. sake* and *M. caribbica* isolates for the investigation of the headspace oxygen level during the shelf-life of pineapple, since they were previously isolated from spoiled commercial fresh-cut pineapple. Ibrahim et al. (2017) identified *Fusarium proliferatum, F. verticillioides, F. sacchari* and *Fusarium* sp. in diseased pineapple tissues. Some of the *Fusarium* sp. isolates appeared to be phylogenetically related to *F. circinatum.* Recently, Lima et al. (2019) found that in tropical fruit-based ice creams (including pineapple based) the predominant species were *Candida intermedia, Torulaspora delbrueckii, Candida parapsilosis, Clavispora lusitaniae, Saccharomyces cerevisiae* and *Pichia kudriavzevii.* It is obvious that there are differences in fungal composition of pineapple samples among various studies. As it was mentioned above, these differences could be attributed to various pre- and post-harvest environmental factors, agricultural practices, but also plant genotype. Additionally, there is a low species diversity in previous studies which may be related to the bias of culture-dependent techniques (as discussed above) or the limitations of conventional genomic methods (Subasinghe et al., 2019).

In the light of the available scientific literature and by comparing the findings of the aforementioned studies with the data obtained in this work, it is evident that pineapple may harbour a rather diverse fungal microbiota. However, the present study provides additional information on how such variable microbiota may behave as function of the applied storage conditions. Indeed, a thorough observation at batch level allows for various conclusions to be

drawn regarding the impact of storage temperature and time. Specifically, the batches P1 and P2 were characterized by the great dominance of *F. circinatum*. Multiple species of *Fusarium* are associated with fruit rot and leaf spot diseases of fruits and especially pineapple (Jacobs et al., 2010). However, at the later stages of spoilage in batch P1, *P. fermentans* prevailed at 4 ^{o}C, while *C. intermedia* prevailed at 8 and 12 ^{o}C. Both *P. fermentans* and *C. intermedia* have been studied with great potential in the biological control of phytopathogenic molds. Rosa-Magri et al. (2011) identified *C. intermedia* as one of the yeasts with biocontrol activity against *Colletotrichum sublineolum* and *C. graminicola*. Giobbe et al. (2007) investigated the dual nature of a strain of *P. fermentans* which controls brown rot on apple fruit, but becomes a destructive pathogen when applied to peach fruit. Consequently, it could be postulated that these two yeasts could possibly play a competitive role in supressing the growth of *F. circinatum* and depending on the temperature, one of the two closely related yeasts is able to dominate. The same trend was observed in the batch P2, but *C. argentea* prevailed finally at 4 ^{o}C. On the contrary, when *C. argentea* was present in great dominance (batch P3) at first place, there was no significant impact of temperature and time. Nonetheless, the conclusion is differentiated when *H. uvarum* was initially present together with *C. argentea* as the second most dominant species (batch P4). These two closely related species seem to be affected by the temperature and time in an opposite way, but to a lesser extent. Consequently, the progress of spoilage appears to depend firstly on the initial composition and secondly on the effect of temperature and time.

With regard to the bacterial microbiota of pineapple, no significant compositional differences were observed among the different pineapple batches and applied storage temperatures. The bacterial populations seem to be subdominant compared to the fungi present to the pineapple. These bacterial communities consisted mainly of *Pseudomonas* spp. which was the most prevalent genus in pineapple samples. Although the pH of pineapple does not allow the growth of most bacteria, the members of this genus may come from the field. Indeed, *Pseudomonas* species have been reported as important plant pathogens in fruits and vegetables (Bophela et al., 2019). *Pseudomonas viridiflava*, a well-known multi-host plant pathogen (Sarris et al., 2012), was the most abundant *Pseudomonas* species in the present study. Members of the Enterobacteriaceae family, as well as LAB, were also present as part of the pineapple microbiota in lower or even similar abundances to *P. viridiflava*. Enterobacteriaceae

may originate from field sources including soil, insects, irrigation water, wastes, animals, but also from human and equipment contamination during harvesting and processing (Erkmen & Bozoglu, 2016). Three lactic acid bacteria species, namely *Lactobacillus plantarum*, *Lactobacillus rossiae* and *Weissella cibaria* were also identified by Di Cagnio et al. (2010) in pineapple flesh, with the former species being the most abundant. Indeed, various LAB have been reported as spoilage members and occasionally as bacterial inhibitors in fruits (Pothakos et al., 2014; Saraoui et al., 2016).

2.5. Conclusions

The findings of the present study contribute to the understanding of the fungal and bacterial communities associated to fresh-cut RTE pineapple. It is the first time that the impact of storage temperature and time on microbial diversity is being studied for a fresh tropical fruit product. With regard to fungi, the results demonstrated that the composition of fungal communities may be different and highly variable among different batches of pineapple. The initial microbial composition of this commodity appears to constitute an important factor with regard to the evolution of its spoilage. Depending on the initial prevalent fungal species, the impact of temperature and storage time varies. It is obvious that fresh-cut pineapple products are a very complex (and hard to predict) ecological niche, where the specific spoilage species may have a totally different response to the changes of important environmental factors prevailing during storage, and which are however important in the shelf-life assessment of these RTE fruit. On the other hand, bacterial communities were not significantly affected neither by the product batch nor by the applied storage temperature. With regard to future research perspectives, further and thorough research is necessary in order to unravel how the various environmental factors of pineapple production may drive the initial pineapple microbial composition. In this view, a large-scale analysis of multiple pineapple samples from various production facilities is advisable.

Acknowledgments

We thank the MIGALE bioinformatics platform at INRAE (http://migale.inrae.fr) for providing computational resources and data storage. We also thank the INRAE@BRIDGe platform for carrying out the MiSeq sequencing runs.

Chapter 3

Metagenetic characterization of bacterial communities associated with ready-to-eat leafy vegetables stored at different temperature conditions

Manuscript for publication in preparation

Evanthia Manthou, Gwendoline Coeuret, Stephane Chaillou, George-John E. Nychas

Chapter 3

Abstract

Ready-to-eat (RTE) and fresh-cut vegetables meet the current needs for healthy and easy-to-prepare food. However, raw vegetables are widely known to harbor large and diverse bacterial communities promoting spoilage and reducing their shelf-life. A better understanding of their bacterial community and the impact of various environmental factors on its composition is essential to ensure the production of safe and high-quality fresh-cut produce. Therefore, a metagenetic amplicon approach, based on *gyrB* sequencing, was applied for deciphering the bacterial species communities associated with the spoilage of RTE rocket and baby spinach, and monitoring the changes occurring in their composition during storage at different temperatures. The results indicated that *Pseudomonas* genus was the main spoilage group for both leafy vegetables. Specifically, *Pseudomonas viridiflava* was dominant in most samples of rocket, while a new *Pseudomonas* species, as well as *Pseudomonas fluorescens* and/or *Pseudomonas fragi* were highly abundant in baby spinach. Moreover, a significant variability in bacterial species composition among different batches of each vegetable type was observed. On the other hand, the impact of storage temperature and/or time on bacterial microbiota was not explicitly revealed for baby spinach. Concerning rocket, storage time was the most influential factor resulting in the reduction of *Pseudomonas* species' abundances and the parallel increase of lactic acid bacteria's abundances. The results suggest that a large-scale sampling of RTE produce combined with a broader range of studied temperatures should facilitate the development of effective microbial control strategies and thus, enhance the safety and quality of RTE produce.

3.1. Introduction

Ready-to eat (RTE) and fresh-cut produce commodities satisfy the ever-growing need for consumption of healthy and high quality food products, combined with time saving and convenient food preparation (Conte et al., 2008; Pilone et al., 2017; Tsironi, et al., 2017). Plant food products are known to harbour high populations of diverse microbial communities originating from various sources, including the farm environment and post-harvest handling (Soderqvist et al., 2017; Gu et al., 2018). In the fresh-cut produce production chain, microbial communities may be further enriched during mild processing resulting in faster quality deterioration (Francis et al., 2012; Qadri et al., 2015). Therefore, the microbial contamination is the cause of significant economic losses for the fresh-cut sector by reducing the shelf-life of the products and increasing the risk of potential foodborne illnesses (Rico et al., 2007; Gorni et al., 2015; Giannopoulou et al., 2020).

The surface of vegetable leaves is inhabited by several microorganisms among which bacteria are considered the most numerous colonists (Rastogi et al., 2012; Cao et al., 2017; Soderqvist et al., 2017). Bacteria can be also found within the plant tissues as endophytes entering through the root system or mechanical damages and could be the source of additional microbial dissemination (Rastogi et al., 2012; Soderqvist et al., 2017). The overall bacterial communities associated with raw vegetables consist of pathogenic and spoilage microorganisms. Moreover, specific members are known to act as potential competitors to pathogens colonization and their addition may control significantly the quality and safety of such products (Schuenzel & Harrison, 2002; Cooley et al., 2006; Klerks et al., 2007; Lopez-Velasco et al., 2010; Soderqvist et al., 2017).

Until recently, leaf microbiota has been typically characterized using traditional cultivation methods or molecular culture-based approaches (Jackson et al., 2015). However, the development of culture-independent DNA-based methods enabled the detection of a broader range of bacteria that are present in low abundances or grow slowly and may have escaped identification (Jackson et al., 2013). Next generation sequencing (NGS) technology and bioinformatics analyses have promoted a finer characterization of microbial communities from various environments than previously obtained using culture-based approaches (Kumar et al., 2015; Goodwin et al., 2016; Cao et al., 2017). Several NGS diversity studies have attempted to assess the impact of plant genotype and origin, season, farming practices, as well

as processing methods on the compositions of vegetables' bacterial communities (Hunter et al., 2010; Rastogi et al., 2012; Jackson et al., 2013; Dees et al., 2015; Cao et al., 2017; Darlison et al., 2019; Tatsika et al., 2019). The majority of the results derived by these studies indicated that the aforementioned factors can strongly influence the bacterial communities of such products. Less common are the NGS studies conducted to monitor the impact of storage conditions, and mainly storage temperature, on the bacterial communities established on the leaves (Lopez-Velasco et al., 2011; Soderqvist et al., 2017; Gu et al., 2018; Rosberg et al., 2020). The main conclusion of these limited studies was that distinct temperatures lead to accordingly different composition and abundance of the leafy vegetable microbiota. However, there is still a limited understanding of the diversity of produce-associated bacterial communities, the way various factors influence and shape these communities, and the distributions of individual taxa across produce types (Leff & Fierer, 2013).

The objective of this study was to characterize the bacterial communities associated with the spoilage of RTE rocket and baby spinach during storage at refrigeration and abusive temperatures. So far, few studies investigating the microbial communities of raw leafy vegetables have focused on the final RTE product (Jackson et al., 2013; Leff & Fierer, 2013; Soderqvist et al., 2017; Tatsika et al., 2019). Among these studies, only Jackson et al. (2013) and Soderqvist et al. (2017) analysed both epiphytic and endophytic communities of RTE baby spinach, and this is the approach that has been also followed in the present work. The innovation of the present study lies in the applied metagenetic amplicon approach based on *gyrB* sequencing. This genetic marker has been previously used in bacterial characterization of meat and seafood microbiota (Poirier et al., 2018; Poirier et al., 2020; Zagdoun et al., 2020) and recently in RTE pineapple (Manthou et al., 2020). The main advantages of the latter approach have been (i) the achievement of more accurate (at intra-species level) taxonomic assignment than classical V3-V4 region of the16S rRNA (usually family- or genus-level), and (ii) the avoidance of contamination by chloroplastic 16S rDNA ,which is very common in 16S-based metagenetic analysis from plant materials. Therefore, the main goal of this study was the investigation and characterization of the bacterial species communities associated with the spoilage of RTE leafy vegetables and the changes occurring in their composition during storage at different temperatures.

3.2. Materials and Methods

3.2.1. Leafy vegetable samples and storage conditions

Fresh-cut and RTE rocket and baby spinach salads, already packed in sealed plastic bags (each one containing 125 g of product) were supplied by a local manufacturer in Athens. All RTE produce was provided by the same manufacturer, except for one batch of rocket salad which supplied by a local supermarket. The salads were transported to the laboratory within 24 h from production and stored in their original package at three different constant temperatures (4, 8 and 12 ^{o}C), as well as at dynamic temperature conditions (8 h at 4 ^{o}C, 8 h at 8 °C and 8 h at 12 ^{o}C). One batch of rocket salad was received and stored few hours after its production, thus corresponding to a more fresh state. The storage took place in high precision (±0.5) programmable incubators (MIR-153, Sanyo Electric Co., Osaka, Japan), while the temperature was recorded every 15 min using electronic temperature devices (COX TRACER®, Cox Technologies Inc., Belmont, NC, USA). On the day of arrival (time-zero) at the laboratory and also at 38, 72, 134 and 230 h, samples from four independent batches of each vegetable were analyzed. The storage at 12 ^{o}C ended at 134 h compared to all the other temperatures that was extended up to 230 h.

3.2.2. Microbiological analysis, pH and gas composition measurements

Prior to microbiological analysis, the O_2/CO_2 composition inside the packages was measured using a headspace gas analyzer (CheckMate 9900, PBI Dansensor, Denmark). Subsequently, a 25 g-portion of each salad was aseptically transferred into a sterile Stomacher bag (Seward Medical, London, UK), diluted with 225 ml of quarter-strength Ringer's solution (Lab M Limited, Lanchashire, UK) and placed in a homogenization device for 60 sec (Lab Blender 400, Seward Medical). Appropriate serial decimal dilutions were prepared and the surface plating technique was used for the determination of total mesophilic microbial populations (total viable count, TVC) and *Pseudomonas* spp. populations. The TVC was determined on tryptic glycose yeast agar (Plate Count Agar, Biolife, Milan, Italy) after incubation of plates at 25 ^{o}C for 72 h. *Pseudomonas* spp. were determined using the same technique on pseudomonas agar base with selective supplement cephalothin-fucidin-cetrimide (CFC, Lab M Limited), after incubation of plates at 25 ^{o}C for 48 h. The results were expressed as the average (± standard deviation, n=4) log colony forming units per gram (log CFU/g) of

vegetable salad. Additionally to microbiological analysis, the pH values of vegetable samples were measured with a digital pH meter (RL150, Russell pH Cork, Ireland) with a glass electrode (Metrohm AG, Herisau, Switzerland).

3.2.3. Preparation of microbial pellets and DNA extraction of the vegetable microbiota

A 10g-portion of each vegetable sample was transferred in filter Stomacher bag (Interscience, St-Nom, France) containing 20 ml of Ringe's solution (Lab M Limited) and homogenized for 60 sec in the stomacher device. The homogenate was collected in 50-ml tubes (SARSTEDT AG & Co. KG, Germany) and centrifuged (Heraeus Multifuge 1S-R, Thermo Electron Co.) at 8000 × g for 20 min at 4 ^{o}C. The obtained bacterial pellet was washed in 20 ml of distilled-dionized water and centrifuged once more under the same conditions. An extra washing step followed, adding 1.7 ml of sterile ultrapure water, the diluted pellet was transferred in 2-ml Eppendorf tubes (SARSTEDT AG & Co. KG), and the bacterial pellet collected after centrifugation (Heraeus Fresco 21, Thermo Scientific) at 17000 x g for 10 min at 4 ^{o}C was stored at -80^{o}C until the DNA extraction procedure. Bacterial DNA from all samples was extracted with the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) according to manufacturer's instructions.

3.2.4. Barcoding PCRs and Illumina Miseq sequencing

Two rounds of PCR amplification took place for the construction of amplicon libraries. The first amplification of the 250 bp region of *gyrB* was performed with the primers F64 (5'-MGNCCNGSNATGTAYATHGG -3') and R353 (5'- ACNCCRTGNARDCCDYCNGA -3'). Forward and reverse primers carried the Illumina 5'- CTTTCCCTACACGACGCTCT TCCGATCT-3' and the 5'-GGAGTTCAGACGTGTGCTCTTCCGATCT-3' tails, respectively. The first round of PCR was performed with the high-fidelity AccuPrime *Taq* DNA polymerase system (Invitrogen, Carlsbad, USA), 20 μM final primer concentration and 5 μL of microbial DNA. Amplification of *gyrB* was performed with the following cycling conditions: 94 ^{o}C for 2 min, followed by 35 cycles of amplification at 94 ^{o}C (30 sec), 55 ^{o}C (60 sec), and 68 ^{o}C (90 sec), with a final extension step of 10 min at 68 ^{o}C. The amplicon size, quality, and quantity of the amplified DNA were checked on a DNA1000 chip (Agilent

Technologies, Paris, France). The second Miseq PCR and the Miseq sequencing were conducted as described in Poirier et al. (2018). Raw read sequences were deposited at the Sequence Read Archive under the Bioproject number PRJNA694950 and the accession numbers SAMN17579174 to SAMN17579262.

3.2.5. Sequencing data analysis

3.2.5.1. Quality filtering and taxonomic assignment of Operational Taxonomic Units (OTUs)

The quality control, definition of OTUs and taxonomic assignment were performed using FROGS (Find Rapidly OTU with Galaxy Solution) pipeline (Escudie et al., 2017). Briefly, the quality-filtered paired-end sequences were merged into contigs with VSEARCH v2.15.0 (Rognes et al., 2016) using the maximum of 10% mismatch in the overlapped region. The minimum and maximum amplicon size was set at 150 and 350 bp, respectively. Merged amplicon sequences were dereplicated and clustered using SWARM v3.0.0 (Mahe et al., 2015) with a distance threshold of three. Chimeras were removed with VSEARCH v2.15.0. The resulting sequences were filtered for spurious OTUs by keeping only those with at least 0.01% of relative abundance within the whole dataset (Auer et al., 2017). Taxonomic assignment of OTUs was performed using *gyrB_03_2019* as reference databank (Poirier et al., 2018) and the Blastn+ algorithm (Camacho et al., 2009).

3.2.5.2. Alpha and beta diversity

Bacterial diversity was analysed using the R package Phyloseq with standard or custom Phyloseq command lines (McMurdie &Holmes, 2013). The median sequencing depth of all samples was used as a normalization step for OTU abundance, before the analyses of alpha and beta diversity.

3.3. Results

3.3.1. Effect of storage temperature on bacterial growth

The initial bacterial load of both vegetables was high (Figure 3.1). The initial level of TVC (mean ± standard deviation, n=4) was 7.31 ± 0.48 log CFU/g for rocket and 6.96 ± 0.25 log CFU/g for baby spinach. At the final stage of storage, the TVC in rocket reached an average

level (± standard deviation) of 7.79 (± 0.26), 8.49 (± 0.22), 8.25 (± 0.33) and 7.85 (± 0.68) log CFU/g at 4 ºC, 8 ºC, 12 ºC and dynamic conditions, respectively. In baby spinach, the final TVC was similar for the majority of temperature conditions. However, the TVC throughout storage was higher in baby spinach, with larger differences being observed at 4 °C. The *Pseudomonas* spp. populations were close to the TVC with a maximum difference of 0.5 log throughout storage. According to preliminary analysis where additional microbial groups (enterobacteria, lactic acid bacteria) were determined, *Pseudomonas* genus is the main component of the leafy vegetable microbiota. The same trend mentioned for TVC was observed for *Pseudomonas* spp., since their growth was also higher in baby spinach compared to rocket throughout storage. As expected, the recorded bacterial growth was faster with increasing storage temperature.

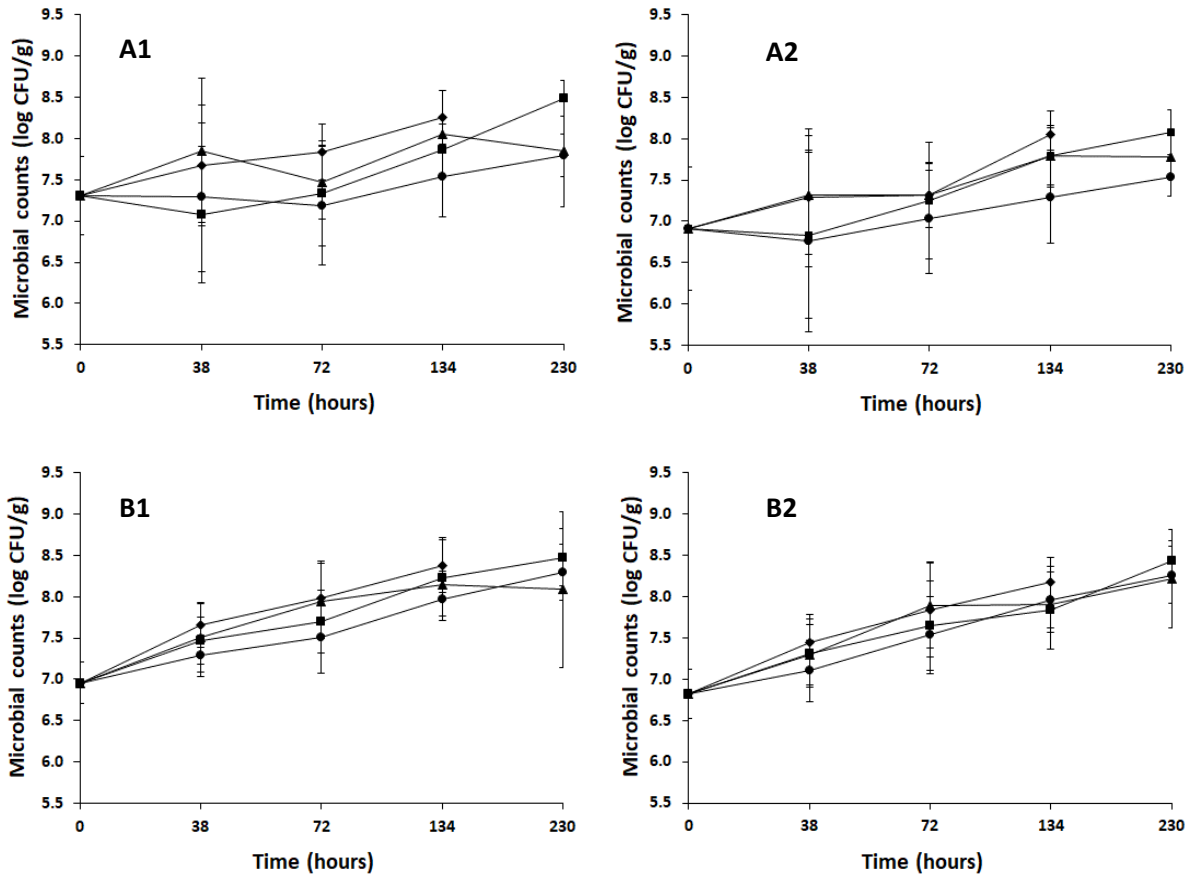


Figure 3.1. Total viable counts (TVC) (A1) and *Pseudomonas* spp. populations (A2) in ready-to-eat (RTE) rocket, and TVC (B1) and *Pseudomonas* spp. populations (B2) in RTE baby

spinach during storage at 4 ºC (●), 8 ºC (■), 12 ºC (◆), and dynamic temperature conditions (▲). The microbial populations are expressed as means ± standard deviations (n=4).

However, the impact of temperature on bacterial growth was not notably large. It is also interesting to note that the variability of bacterial counts among the different batches of rocket, was higher than that of baby spinach. One of the rocket batches supplied by a different manufacturer (batch R1) had higher initial TVC and *Pseudomonas* spp. counts compared to the other batches, with a recorded difference as high as 1.5 log units in some cases. Moreover, a second batch of rocket was stored in a more fresh state, having considerably lower bacterial loads at the initial stage of storage.

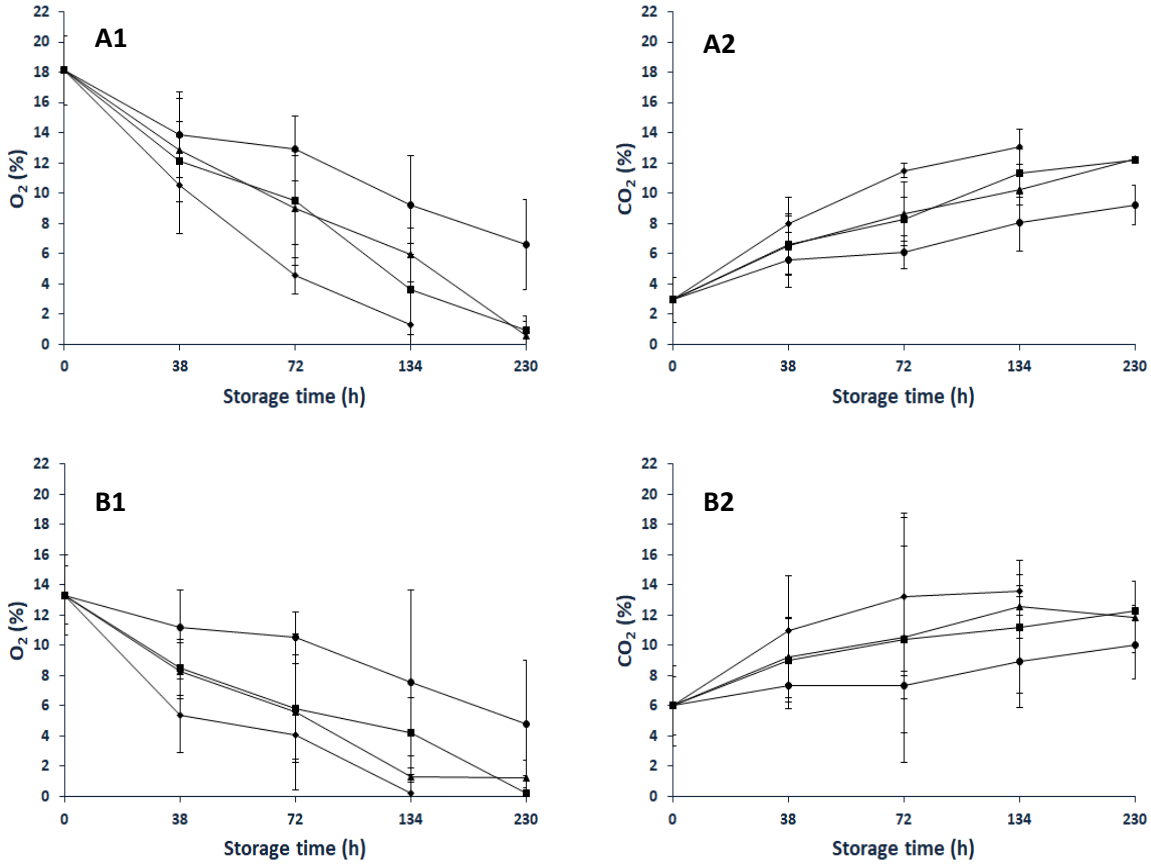


Figure 3.2. O_2 (A1) and CO_2 (A2) compositions in ready-to-eat (RTE) rocket, O_2 (B1) and CO_2 (B2) compositions in RTE baby spinach during storage at 4 ºC (●), 8 ºC (■), 12 ºC (◆), and dynamic temperature conditions (▲) and under passive modified atmosphere packaging.

For rocket, the results concern only the batches R2, R3 and R4. The results are expressed as mean ± standard deviation (n=4 or n=3).

With regard to gas composition inside the sealed packages (Figure 3.2), the initial O_2 level composition in baby spinach was lower than those of rocket, while the initial level of CO_2 was higher. The final O_2 compositions in rocket reached values quite close to those of baby spinach, while the final CO_2 percentages were similar for both vegetables. The storage temperature influenced considerably and similarly the respiration rate for the two vegetables. The consumption and production of O_2 and CO_2 respectively, was higher at increasing temperatures, while in dynamic conditions resembled that of 8 ^oC. It should be also mentioned that in the case of rocket the gas composition of the three out of four batches is presented in Figure 3.2. As it was mentioned above, the batch R1 was supplied by a different manufacturer. The initial O_2 composition was 15.80 % and reached 11.70, 13.40, 15.00 and 11.90 % for 4, 8, 12 ^oC and dynamic conditions, respectively. The final CO_2 did not exceed 1.5 % throughout storage and for all temperatures, indicating that the packaging film has low gas permeability compared to those used for the other batches. In general, there were great variations in gas composition during storage even among different batches originating from the same manufacturer. Moreover, the initial pH of baby spinach was higher than that of rocket and reached also higher values during storage at the same temperatures (Table 3.1). Overall, no considerable differences in the pH values of rocket were observed among the different storage temperatures and during storage. Contrarily, the increase of pH during the storage of baby spinach was lower at 4 °C compared to the higher temperatures.

Table 3.1. The pH values of ready-to-eat rocket and baby spinach during storage at different temperatures.

		pH	
Product	**Storage temperature**	**Initial pH**	**Final pH**
Rocket	4oC	6.34 ± 0.13	6.48 ± 0.26
	8oC		6.86 ± 0.18
	12oC		6.46 ± 0.25
	Dynamic		6.58 ± 0.23
Baby spinach	4oC	6.49 ± 0.07	6.79 ± 0.13
	8oC		7.01 ± 0.14
	12oC		6.75 ± 0.12
	Dynamic		6.96 ± 0.03

3.3.2. Bacterial alpha-diversity

The rarefraction curves (Figure 3.3) performed on quality-filtered reads indicated that sequencing depth was sufficient for all samples tested. Across all samples, a total of 373 and 370 *gyrB* OTUs were detected for rocket and baby spinach, respectively. As shown in Figure 3.4, the bacterial diversity between rocket and baby spinach using data merged at the species-level was compared. Unlike species richness showing similar range of species in both types of product, Shannon indice was significantly (ANOVA, $p<0.05$) different for the two vegetables, indicating that baby spinach bacterial evenness was higher than that of rocket.

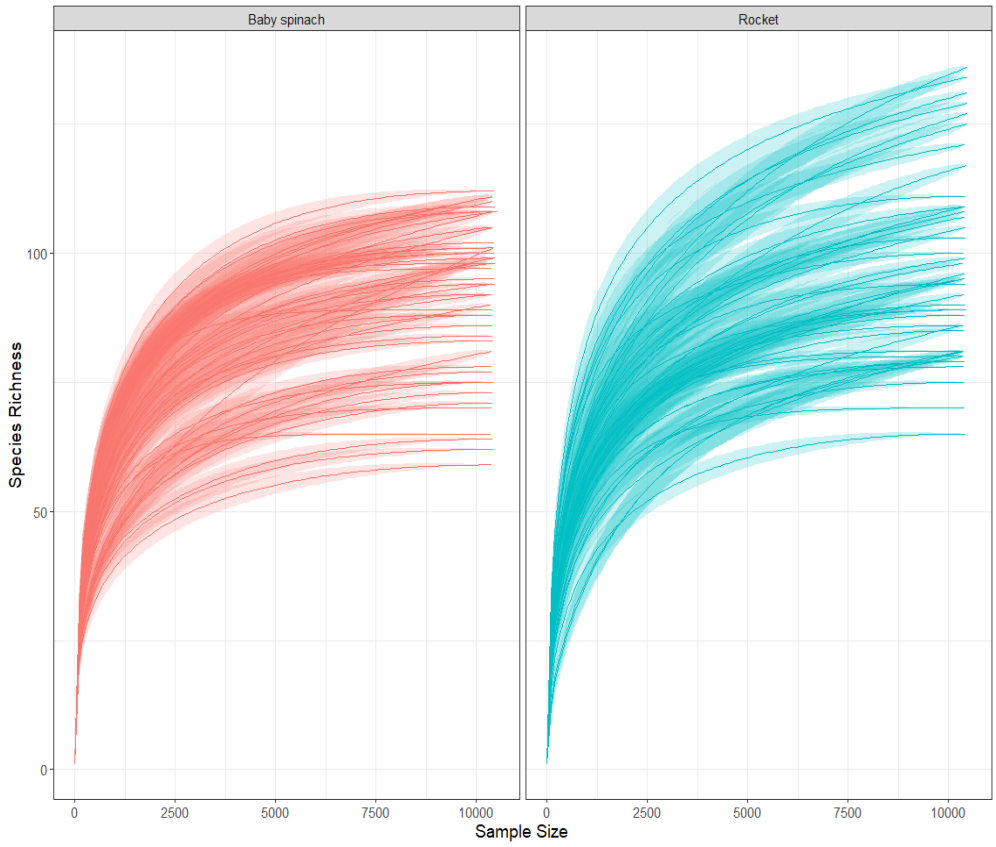


Figure 3.3 Rarefaction curves obtained from *gyrB* amplicon sequencing of the ready-to-eat baby spinach and rocket communities. The x-axis represents the sequencing depth in number of reads while the y-axis represents an estimation of the species richness detected.

To better illustrate this finding, a composition plot was produced (Figure 3.5) comparing the relative abundances of bacterial species according to Bray-Curtis hierarchical clustering of rocket and baby spinach samples. This analysis confirmed the presence of *Pseudomonas* spp. as the main component of vegetable leaf microbiota. Exceptions were observed for some samples of rocket originating from batch R2 that showed great prevalence of *Xanthomonas campestris*. All the samples of batch R3 and some samples of batch R2 were clustered together with baby spinach. Interestingly, there was a clear difference between the two vegetables. *Pseudomonas viridiflava* reached higher abundances, but also was dominant among *Pseudomonas* group and in some cases among all bacterial species for most of the rocket samples, while *Pseudomonas RIT357* was present in very low abundances. However, almost all samples of batch R3 showed a great dominance of *P. fluorescens*, but had very low abundances of *P. viridiflava* compared to the other rocket batches.

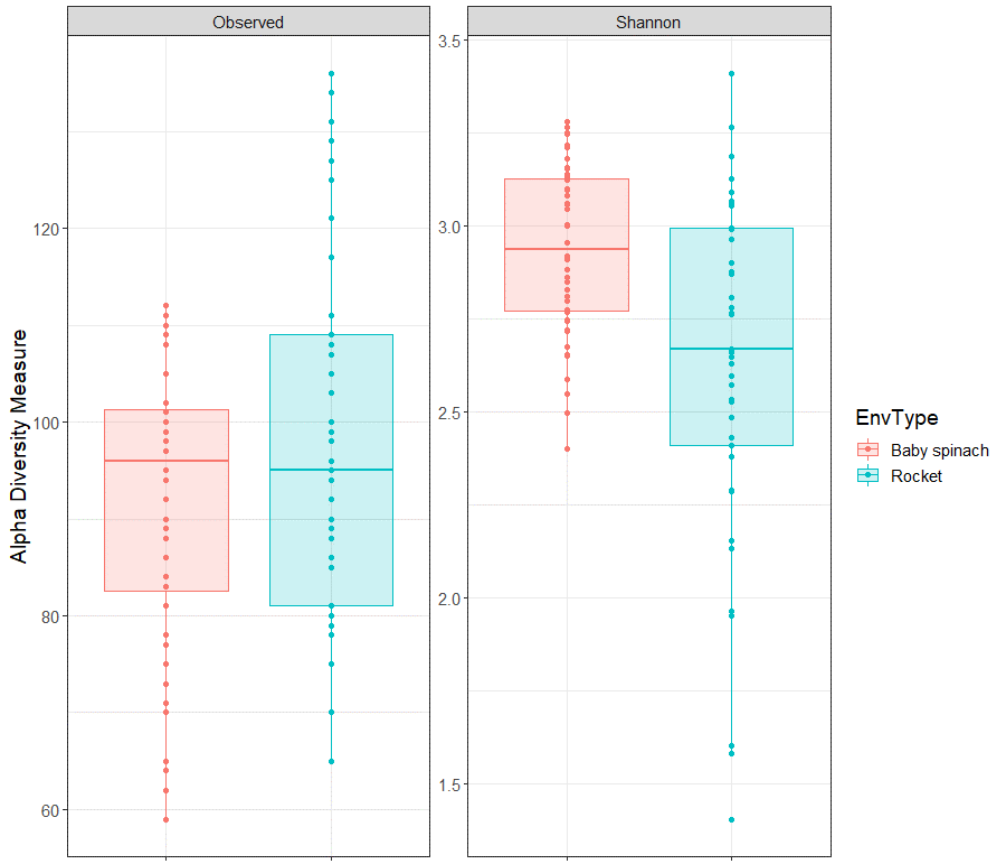


Figure 3.4. Comparison of ready-to-eat (RTE) rocket and baby spinach bacterial diversity. The box plots show the number of species (Observed) and the Shannon indice as measures of species diversity between RTE rocket and baby spinach communities. The boxes represent the interquartile range between the first and third quartiles and the vertical line inside the boxes is the median obtained from the samples analysed per vegetable.

On the contrary, in the case of baby spinach, *P. RIT357* was present in considerably higher abundances compared to *P. viridiflava*. The former was the dominant species or prevailed together with *Pseudomonas fragi* or *Pseudomonas fluorescens* over the other Pseudomonas species. Moreover, baby spinach samples showed higher abundances of *Janthinobacterium S3-2* and in some samples higher abundances of *Janthinobacterium lividum* compared to rocket.

It becomes evident that batches influenced the bacterial communities' composition, and that the batches of rocket had larger compositional differences than baby spinach batches. Specifically, although batch R4 was supplied by the same manufacturer as batches R2 and R3,

its composition was closer to that of batch R1. Even though the rocket batches R1 and R4 were clustered closer to each other, there were obvious differences among them. Specifically, the batch R1 had higher abundances of *Pseudomonas syringae* and *Lactococcus piscium* compared to the other batches.

3.3.3. Effect of storage temperature and time on bacterial diversity

The effect of storage time and temperature was studied separately for each vegetable salad. In the case of rocket, the number of species (observed) was not significantly different among storage times, but Shannon indice (ANOVA, $p<0.05$) increased with storage time (Figure 3.6). This means that there was more equilibrium in the abundance of dominant and subdominant species over longer storage time. The temperature seems that did not influence significantly the species diversity of rocket (data not shown). On the contrary, the impact of temperature (and not of storage time) on baby spinach diversity was significant. Again, the number of species (observed) was not significantly different among the different temperatures, but Shannon indice (ANOVA, $p<0.05$) followed an upward course with temperature increase (Figure 3.7). The samples stored at dynamic conditions were between the samples stored at 8 and 12 ^{o}C, as it was expected, while baby spinach samples stored at 12 ^{o}C had a more diverse community.

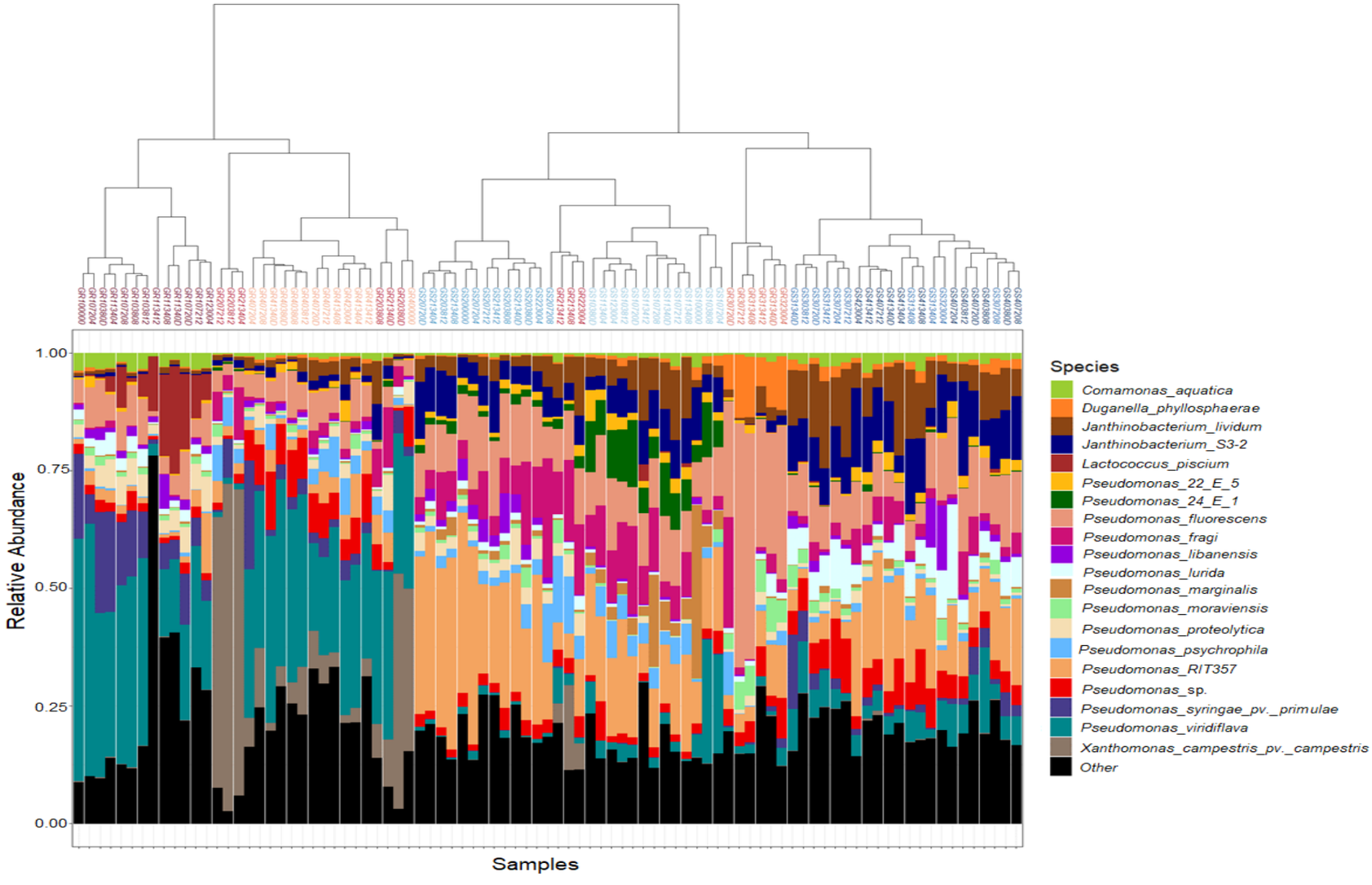


Figure 3.5. Composition plot showing the relative abundances of the 19 main species found in vegetable samples. On the top: hierarchical clustering of batches samples according to Bray-Curtis distance and ward algorithm (different shades of red and blue represent the four batches of ready-to-eat rocket and baby spinach, respectively). The two first characters of each sample are indicating the type of vegetable (e.g. GR and GS for rocket and baby spinach, respectively). The number following these two letters depicts the batch number.

When the impact of temperature and storage time on RTE baby spinach and rocket composition was studied, there was no discrete clustering among samples of different temperatures and storage times for each one of the two vegetables (data not shown). According to Figure 3.8, the batches had a strong influence on the type of community, indicating that there is a little chance to assess the impact of both factors on the overall composition of both vegetables.

A deeper analysis at the batch level did not reveal a clear effect of storage time and/or temperature on bacterial diversity for baby spinach. At Batch BS1, *Pseudomonas marginalis* had a high abundance at time-zero, but it was present in very low abundances for almost all the temperatures and without gradual change for the following storage stages (Appendix I, Figure S.3.1). Moreover, *J. lividum* showed a slight higher abundance during storage and the change was sharper at 8, 12 ºC and even dynamic conditions. At Batch BS2, *P. RIT357* reached higher abundances at 4 ºC compared to the most but not all samples stored at the higher temperatures (Appendix I, Figure S.3.2). For the two other batches of baby spinach, it was difficult to reveal any impact of time and temperature (data not shown). In the case of rocket, on the other hand, the storage time was the most influential factor for some species of Batch R1. Data shown in Figure 3.9 indicate that rocket samples of Batch R1 are separated in two branches, one containing most samples stored under 4 °C and 8 °C, whereas, the second branch those stored under 12 °C or dynamic conditions. Furthermore, storage time affected this clustering as samples from early storage for 12 °C and dynamic are located in the left branch of the tree, whereas samples of 4 °C and 8 °C corresponding to long storage time are clustered on the right branch. More specifically, *P. viridiflava* lost ground in the storage time course for each one of the four temperatures, while *Carnobacterium inhibens* and *Lactococcus piscium* (lactic acid bacteria) became more abundant. At the first stages of storage (0, 38, 72 h), *P. viridiflava* was the most abundant species; its abundance reduced during storage but it remained the most prevalent species in samples stored at 4 ºC. On the other hand, although *C. inhibens* showed small increase of abundance in all cases, *L. piscium* became prevalent at 8 ºC and dynamic conditions. Moreover, the abundance of *P. syringae* was also reduced during storage, but this observation was clearer at 8 and 12 °C and at dynamic conditions. Concerning Batch R2 (Figure 3.10), *Xanthomonas campestris* showed reduction of abundance over longer storage time mainly at 12 °C, but also at 4 °C and dynamic conditions. In Batch R3, there were

missing a lot of samples for solid conclusions to be drawn, while Batch R4 did not show any significant changes in terms of time and temperature, regardless of the high prevalence of *P. viridiflava* and the closest composition similarity with Batch R1 compared to the other batches (data not shown).

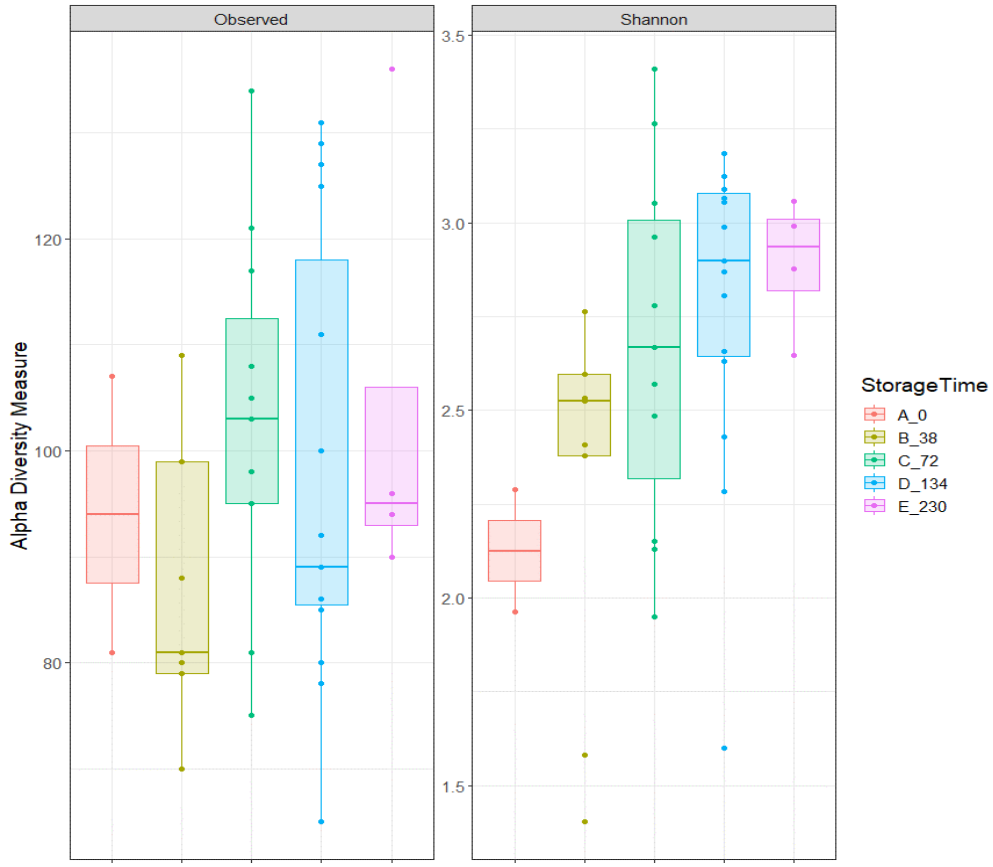


Figure 3.6. Bacterial diversity in ready-to-eat rocket. The box plot shows the number of species (Observed) and the Shannon indice as measures of species diversity in samples of different storage time. The boxes represent the interquartile range between the first and third quartiles and the vertical line inside the boxes is the median obtained from the samples analysed per condition. A_0, B_38, C_72, D_134 and E_230 represent 0, 38, 72, 134 and 230 h of storage, respectively.

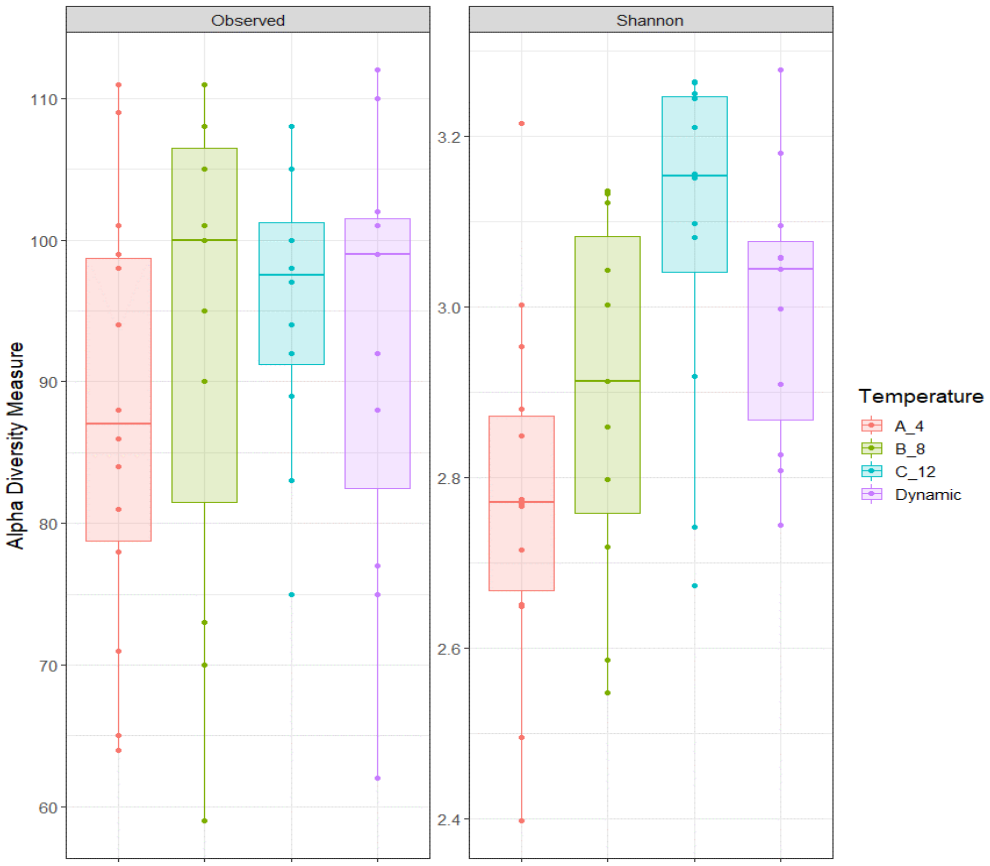


Figure 3.7. Bacterial diversity in ready-to-eat baby spinach. The box plot shows the number of species (Observed) and the Shannon indice as measures of species diversity in samples of different temperature. The boxes represent the interquartile range between the first and third quartiles and the vertical line inside the boxes is the median obtained from the samples analyzed per condition. A_4, B_8, C_12 and Dynamic represent 4, 8, 12 ºC and dynamic temperature conditions, respectively.

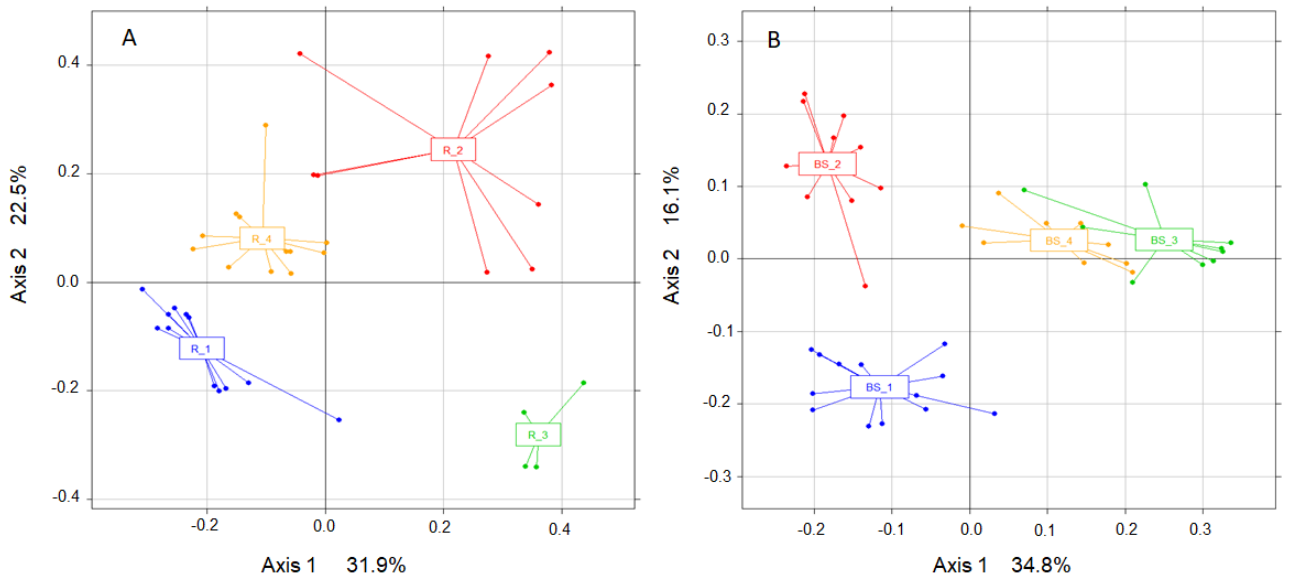


Figure 3.8. Multidimensional Scaling (MDS) based on Bray-Curtis distances among bacterial communities of the four batches of ready-to-eat A) rocket and B) baby spinach. R_1, R_2, R_3 and R_4 or correspond to Batch 1, Batch 2, Batch 3 and Batch 4 of rocket and BS_1, BS_2, BS_3 and BS_4 correspond to Batch 1, Batch 2, Batch 3 and Batch 4 of baby spinach.

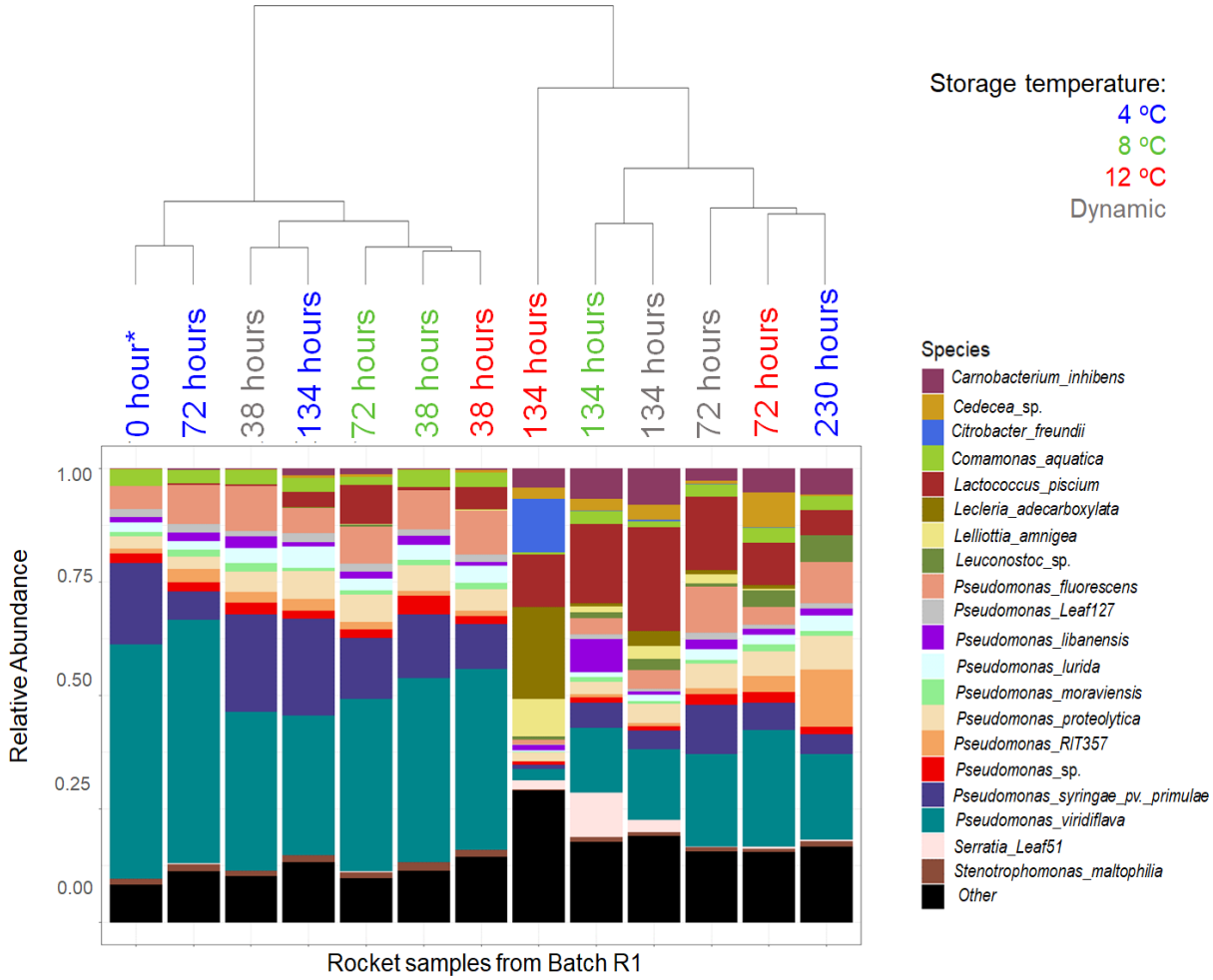


Figure 3.9. Impact of storage time and temperature on bacterial species composition of ready-to-eat rocket samples from Batch R1. Samples are ordered from left to right according to Bray-Curtis distance. Samples are colored according to storage temperature in blue (4 °C), green (8 °C), red (12 °C) and Dynamic (grey).

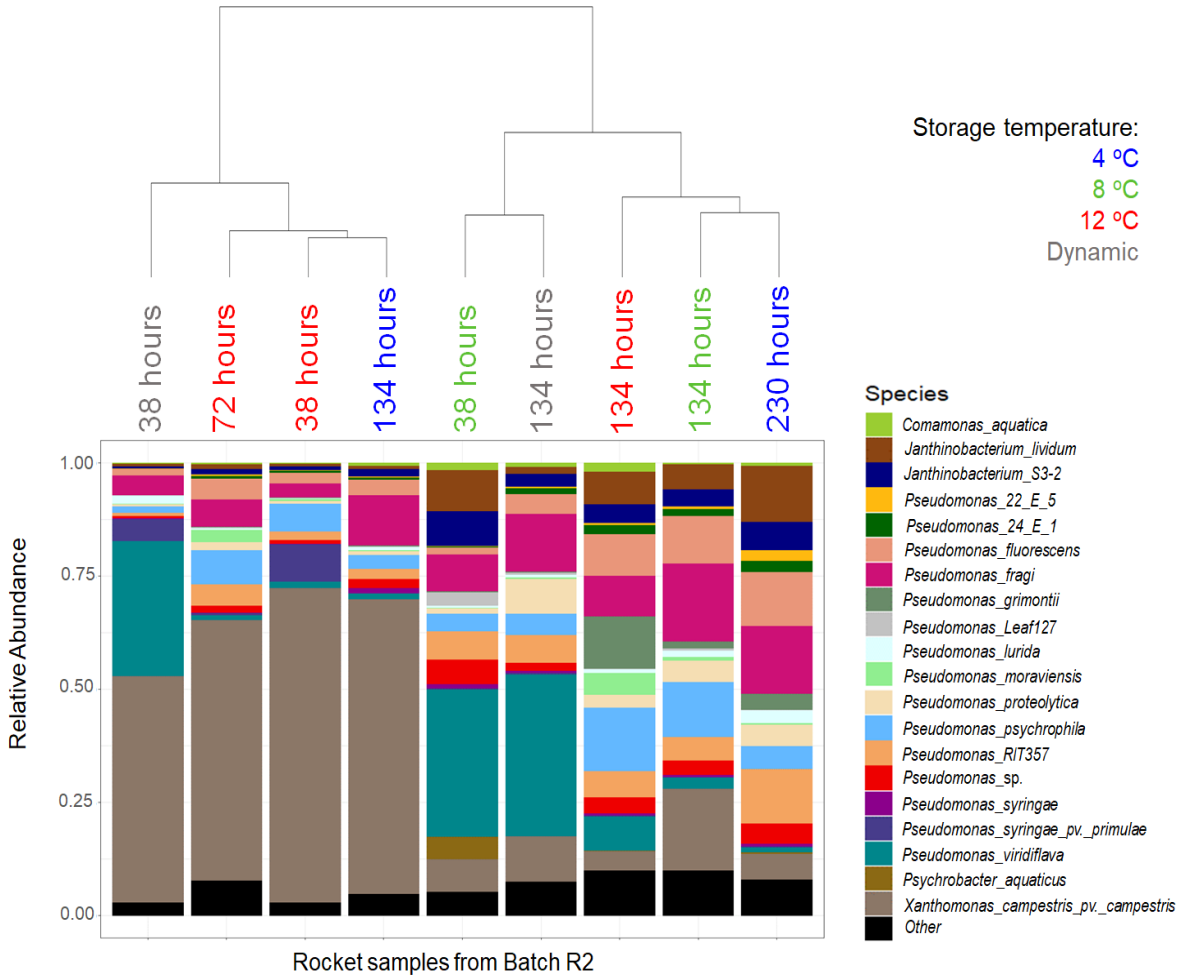


Figure 3.10. Impact of storage time and temperature on bacterial species composition of ready-to-eat rocket samples from Batch R2. Samples are ordered from left to right according to Bray-Curtis distance. Samples are colored according to storage temperature in blue (4 °C), green (8 °C), red (12 °C) and Dynamic (grey).

3.4. Discussion

In the present study, *gyrB* amplicon sequencing was applied to characterize the bacterial communities associated with RTE leafy vegetables and examine the impact of different temperatures on the composition of the vegetable microbiota during storage. The results of this study indicated from both the plate counting and metagenetic approach, that the *Pseudomonas* genus was the dominant spoilage group in the majority of rocket samples and all samples of baby spinach. Indeed, *Pseudomonas* spp. have been previously reported as the most prevalent bacteria in raw leafy salad vegetables, including RTE rocket and baby spinach (Rastogi et al., 2012; Jackson et al., 2013; Soderqvist et al., 2017; Rosberg et al., 2020). Members of this

genus are widely distributed in the nature and are highly capable of competing with other bacteria in food and food production environments (Hibbing et al., 2010). Some species are recognised as potential plant pathogens and spoilage causative agents since they have the ability to produce pectinolytic enzymes that cause soft rot and fleshy vegetables (Tatsika et al., 2019). Contrarily, Tatsika et al. (2019) observed that in RTE baby spinach but not rocket, the *Enterobacteriales* order had higher abundances compared to *Pseudomonadales* which is mostly represented by the *Pseudomonas* genus. Also, Jackson et al. (2013) have reported other genera that may be prevalent to various leafy vegetables, such as *Ralstonia, Xanthomonas, Flavobacterium, Stenotrophomonas Serratia* and *Erwinia*.

Despite the increasing number of metagenetic studies dealing with fresh produce, almost all of them remain at the taxonomic level of family or genus. Although the application of 16S rDNA amplicon sequencing is one of the most popular techniques for the characterization of food microbiota, it lacks the necessary resolution required to provide species-level/strain-level identification (Cao et al., 2017). However, spoilage development is a complex biological phenomenon, which can be species- or even strain-specific. Moreover, certain microbial species may be influenced differently by a given set of storage conditions, driving unpredictably the time and type of spoilage (Kergourlay et al., 2015; Poirier et al., 2018). Therefore, a deeper understanding of microbial ecology is needed with the primary goal to monitor microbiological food spoilage and safety. The present study succeeded, for the first time, in reaching a species-level characterization of the bacterial community of the studied RTE vegetables, using *gyrB* genetic marker.

According to diversity indices, although the species richness was not significantly different between the two vegetables, baby spinach showed higher species evenness than rocket. The bacterial community of baby spinach has been previously reported as more diverse, but in terms of OTU richness (Tatsika et al., 2019). As far as the differences in species composition are concerned, high abundance and occasionally dominance of *P. viridiflava* in RTE rocket were observed. This species is a soft-rotting spoilage bacterium capable of acting at and below 4 ^{o}C and it is reported as prevalent in decayed vegetables (Saranraj et al., 2012; Sarris et al., 2012). On the contrary, a new species, namely *P. RIT357*, which is an isolate from trees of the *Salix* genus, was highly abundant or even dominant in baby spinach. *P. fluorescens* and/or *P. fragi* were also highly abundant in baby spinach and have long been known to be

responsible for chilled food spoilage (Franzetti & Scarpellini, 2007). Moreover, *J. S3-2*, which is currently published as *Janthinobacterium psychrotolerans*, as well as *J. lividum* showed higher relative abundances in most samples of baby spinach compared to rocket. The former species was initially isolated from a frozen freshwater pond and has also the ability to grow at very low (even -3 ^{o}C) temperatures (Gong et al., 2017). The latter species (i.e. *J. lividum*) has been previously found in romaine and green leaf lettuce, as well as RTE escarole and red chicory, and is commonly isolated from soils and water of rivers, lakes and springs (Pantanella et al., 2006; Jackson et al., 2013; Gaglio et al., 2019). Several factors contribute to the community patterns and it is often difficult to determine which of them are responsible for driving the divergence between the bacterial communities on different produce types. Leaf bacterial communities are known to differ across plant species due to variations in metabolites, physical characteristics, and symbiotic interactions with the host (Redford et al., 2010; Leff & Fierer, 2013). However, the findings of this study indicated that the different batches differed slightly or largely in bacterial community composition, even when originating from the same supplier, an observation which was more intense in rocket (Figure 3.8). Differences in growing conditions, handling, transport procedures, and minimal processing (e.g. washing or cutting) can also shape the microbial community composition and may be responsible for these variations among batches (Leff & Fierer, 2013).

The present study also provided additional information on how different microbiota may behave as affected by storage conditions. The role of storage at low temperatures is known to reduce bacterial growth. Indeed, the growth of total mesophiles and *Pseudomonas* spp. populations was faster at higher temperatures and this is also reported by other researchers (Conte et al., 2008; Gu et al., 2018; Georgopoulou et al., 2020). Previous diversity studies have also revealed that storage of RTE spinach or rocket at different temperatures and prolonged time influenced the bacterial diversity and composition at least in higher taxonomic levels (Lopez-Velacso et al., 2010; Lopez-Velacso et al., 2011; Soderqvist et al., 2017, Gu et al., 2018; Rosberg et al., 2020). In the present study, the diversity analysis indicated that the impact of temperature was significantly stronger in baby spinach evenness, while the storage time influenced more the rocket diversity in terms of evenness. However, when the impact of the aforementioned factors on the bacterial community composition of each type of vegetable was investigated, the inter-batch impact proved to be stronger.

The deeper analysis at batch level did not reveal specific and clear trends with regard to the impact of temperature and time on baby spinach communities (Appendix I). On the other hand, the impact of storage time was evident in some batches of rocket, as it was also indicated by the diversity indices (Figure 3.5). The most striking observation was that in Batch R1, where members of *Pseudomonas* genus, mainly *P. viridiflava* but also *P. syringae,* became less abundant during storage. Both species are multi-host plant pathogens of great economical importance (Morris et al., 2007; Sarris et al., 2012). In parallel to *Pseudomonas* species reduction, *L. piscium* and *C. inhibens* showed increase of their abundance during storage. Many studies have now widely studied lactic acid bacteria such as *Carnobacterium, Lactobacillus, Lactococcus* and *Leuconostoc*, for biopreservation and bacterial inhibition in fruits, fermented and raw vegetables. *L. piscium* has been previously isolated from RTE minimally processed vegetable salads and is described either as a bioprotective or spoilage microorganism depending on the strains and the food matrix (Pothakos et al., 2014; Saraoui et al., 2016). Moreover, in the case of Batch R2, *X. campestris* which is an economically important bacterial plant pathogen causing black rot disease (Papaianni et al., 2020), became less abundant throughout storage at almost all temperatures, while various other *Pseudomonas* species showed increased abundances. Various *Pseudomonas* species have been also reported as bacterial antagonists to pathogenic molds, but also bacteria, in fruits and vegetables, depending again on the product and the microbial strain (Sharma et al., 2009).

3.5. Conclusions

The present study provided insights into the bacterial community compositions and dynamics of RTE leafy vegetables as affected by storage conditions. It was the first time that a metagenetic approach allowed for bacterial characterization at the level of species. The results demonstrated that the prevalence of some *Pseudomonas* species in bacterial communities was vegetable-specific. However, product batch was evaluated as the factor with the strongest impact on bacterial communities of RTE baby spinach and rocket. This work also highlighted the difficulty in revealing the impact of important environmental factors on bacterial communities of fresh-cut produce. The impact of storage temperature and time was

not revealed for the baby spinach batches, while for rocket only the impact of storage time was observed in some cases.

Given the high biological variability of plant products, a large-scale sampling of RTE produce is needed in order to allow for improved and integrated analysis, resulting in turn, in deeper knowledge of the diverse bacterial communities on leafy vegetables. The study of the biodiversity of these communities is also related to the need of a more detailed information on the bacterial interactions specifically associated with food spoilage. Moreover, the various environmental factors shaping the microbial communities during storage and before consumption of leafy vegetables should be further investigated. Then, the manipulation of bacterial interactions (e.g., promoting specific desirable inetractions) in tandem with the application of appropriate storage conditions, could bring scientists one step closer to the development of effective control strategies assuring quality and safety of RTE produce.

Acknowledgments

We thank the MIGALE bioinformatics platform at INRAE (http://migale.inrae.fr) for providing computational resources and data storage. We also thank the INRAE@BRIDGe platform for carrying out the MiSeq sequencing runs.

SECTION B

Application of spectroscopy-based technologies coupled with machine learning for the assessment of the microbiological spoilage of fresh-cut produce

Chapter 4

Spectroscopy and multispectral imaging technologies as means of assessing ready-to-eat pineapple quality

Published in Computers and Electronics 175 (2020), 105529 (Appendix II)

'Application of spectroscopic and multispectral imaging technologies on the assessment of ready-to-eat pineapple quality: A performance evaluation study of machine learning models generated from two commercial data analytics tools'.

***Evanthia Manthou**, Sergio-Llaneza Lago, Evaggelos Dagres, Alexandra Lianou, Panagiotis Tsakanikas , Efstathios Panagou, Maria Anastasiadi, Fady Mohareb, George –John E. Nychas*

Chapter 4

Abstract

Recently, rapid, non-invasive analytical methods relying on spectroscopy and hyper/multispectral imaging, are increasingly gaining popularity in food science. Although such instruments offer a promising alternative to the conventional methods, the analysis of generated data demands complex multidisciplinary approaches based on data analytics tools utilization. Therefore, the objective of this work was to: (i) assess the predictive power of different analytical technologies (sensors) coupled with machine learning algorithms in evaluating quality of ready-to eat (RTE) pineapple (*Ananas comosus*), and (ii) explore the potential of The Unscrambler software and the online machine learning ranking platform, SorfML, in predictive model development for the purpose of food quality assessment. Pineapple samples were stored at 4, 8, 12 °C and dynamic temperatures and were subjected to microbiological (total viable count, TVC) and sensory analysis (colour, odour, texture) with parallel acquisition of spectral data. Fourier transform infrared (FTIR), near infrared (NIR), UV/VIS spectroscopies, as well as multispectral imaging (MSI) were used. The results from both analytics tools revealed similar trends. For TVC, almost all the combinations of sensors and partial least squares regression (PLSR) algorithm from both analytics tools reached values of root mean square error (RMSE) of prediction up to 0.63 log CFU/g, as well as the highest coefficient of determination values (R^2). Moreover, linear support vector machines (SVM Linear) combined with almost each one of the sensors reached similar performance. For odour, the model based on FLUO spectral data achieved the highest overall performance, when combined with partial least squares discriminant analysis (PLSDA) in both platforms with accuracy close to 85%, but also with values of sensitivity and specificity above 85%. The SVM Linear and MSI combination achieved similar performance. On the other hand, all models developed for colour and texture showed poor prediction performance. Overall, all the sensors tested with the exception of NIR, constitute promising tools for the microbiological quality evaluation of RTE pineapple.

4.1. Introduction

In the context of tremendous technological change, a growing lack of natural resources, and a continuous evolution of consumers' life-styles and consumption habits across the globe, food industry is challenged to provide safe and qualitative food to consumers. To address the need for efficient, safe and environmentally respectful production, as well as strict communication and connection with the consumers, several approaches have been developed (Nychas et al., 2016). Among these, analytical methods based on vibrational spectroscopy and hyperspectral / multispectral imaging have gained scientists' attention, since they could fulfill the needs of food industry as rapid and efficient methods for assessing food quality, safety and authentication (Papadopoulou et al., 2011; Gromski et al., 2015; Estelles-Lopez et al., 2017).

In contrast with the time-consuming and expensive conventional and molecular-based techniques, the aforementioned approaches constitute a non-destructive and sensible alternative, also suitable for in-, on- and at-line monitoring (Kumar et al., 2014; Nychas et al., 2016; Efenberger-Szmechtyk et al., 2018). Such tools have been reported in the literature as promising tools for quality and safety assessment of different meat products (Ammor et al., 2009; Prieto et al., 2009; Papadopoulou et al., 2011; Panagou et al., 2014; Barbin et al., 2015; Grewal et al., 2015; Ropodi et al., 2015; Trinderup et al., 2015; Ropodi et al., 2017; Fengou et al., 2019a), fish (Fengou et al., 2019b), as well as fruits (Camps & Christen, 2009; Suhandy et al., 2009; Unay et al., 2011; Coldea et al., 2013; Liu et al., 2015) and vegetables (Løkke et al., 2013; Sravan-Kumar et al., 2015; Tsakanikas et al., 2018).

It should be noted that although these instruments/approaches can be considered as efficient, the multivariate nature of the sensor output, is rather complex and usually needs processing and/or dimensionality reduction, before the results can be interpreted (Jollife & Cadima, 2016). Nowadays, in the food sector, a plethora of machine learning approaches has been proposed by different authors in order either to qualitatively or quantitatively predict safety and quality of different foods (Ropodi et al., 2016; den Besten et al., 2018). At the same time, open source platforms are contributing to the improvement of food quality and safety management systems (Nychas et al., 2008; Tenenhaus-Aziza & Ellouze, 2015).

Indeed, the need of computational tools in the area of food science /microbiology has been recognised, due to their capacity to analyse high volumes of heterogeneous data generated from innovative technologies (Roberts & Cozzolino, 2016; Granato et al., 2018; Truong et al.,

2019). This trend is clearly followed by the development of various algorithms including, among others, ordinary least squares (OLS), stepwise linear modelling (SL), principal component analysis (PCA), partial least squares (PLS), support vector machines (SVM), random forests (RFs) and k-nearest neighbours (kNN) which can be applied through free or commercial software with user-friendly and easy-to-use interface. However, choosing the appropriate machine learning approach based on the question that should be addressed, is often challenging and involves a comparative analysis among various algorithms in order to achieve the best possible and realistic performance. This procedure often requires strong statistical and deep interpretation knowledge (Estelles-Lopez et al., 2017).

Even though, the spectroscopy and multispectral imaging techniques have been implemented in a broad range of food products, the application of these technologies to fresh-cut and RTE produce such as pineapple (*Ananas comosus*), is limited, regardless their popularity and market value. Lunadei et al. (2011) evaluated the enzymatic browning of fresh-cut apple slices using multispectral imaging, while a second study of Lunadei et al. (2012) focused on the colour quality of RTE spinach leaves using the same analytical method. Recently, Tsakanikas et al. (2018) studied the microbial quality of RTE green salads (rocket and baby spinach) using different non-invasive sensors based on spectroscopy. As far as RTE pineapple is concerned, only Di Egidio et al. (2009) have studied its shelf-life using vibrational spectroscopy. Therefore, the aims of this work were: (i) to develop mathematic models based on data derived from different analytical instruments to evaluate the sensory and microbiological quality of RTE pineapple, (ii) to compare the models' performance and assess the suitability of different algorithms and analytical technologies for monitoring the various features, and (iii) to explore, the capabilities and the limitations of each data analytical tool.

4.2 Materials and methods

4.2.1. Sample preparation and storage conditions

Fresh-cut and RTE pineapple, packed in PVC trays (each containing 220 g of fruit), was supplied by a local manufacturer in Athens and transported to the laboratory within 24 h from production. The pineapple samples were stored in their original package at three different temperatures, namely 4, 8, 12 ^{o}C, and under dynamic temperature conditions (8 h at 4 ^{o}C, 8 h at 8 ^{o}C and 8 h at 12 ^{o}C) in high precision (±0.5^{o}C) programmable incubators (MIR-153, Sanyo

Electric Co., Osaka, Japan). The incubation temperature was recorded at 15-minutes intervals using electronic temperature devices (COX TRACER®, Cox Technologies Inc., Belmont, NC, USA). The sampling was conducted at regular time intervals, depending on the storage temperature, for a maximum period of 10 days. Specifically, the analyses were carried out every 14 and 10 h ,according to the following sampling time points: 0, 14, 24, 38, 48, 62, 72, 86, 96, 110 h, for the first 5 days and every 24 h until the end of storage. The final time points were 230 h for 4, 8 ^{o}C and the dynamic temperature conditions, and 134 h for storage at 12 ^{o}C.

For each sampling time point, duplicate samples originating from the same temperature conditions but different trays were analyzed. Each sample (tray) was subjected to the following analyses: (i) microbiological analysis and pH measurements; (ii) sensory analysis; (iii) Fourier transform infrared (FTIR) spectroscopic measurements; (iv) near infrared (NIR) spectroscopic measurements; (v) fluorescence (FLUO) spectroscopic measurements; (vi) visible (VIS) spectroscopic measurements; and (vii) multispectral image (MSI) acquisition. Different pineapple parts of the same tray were used for the microbiological analysis to prevent any contamination of the samples during the spectroscopic measurements. Four independent storage experiments were conducted, using four different batches of pineapple. In the case of the fourth experimental replication and only for FLUO and VIS data, the corresponding spectroscopic measurements were carried out every 24 h throughout storage. Moreover, the NIR measurements were performed only for the two pineapple batches, since the sensor was not available for all the experimental period. Consequently, the total number of samples for FTIR and MSI sensors was 424, for FLUO and VIS was 392, while for NIR was 180. However, for TVC model development, the samples were slightly fewer, since there were no microbiological data due to plate contamination. Specifically, the total number of samples for TVC prediction were 417 for FTIR and MSI sensors, 385 for FLUO and VIS, and 177 for NIR.

4.2.2. Microbiological analysis and pH measurements

A 25-g portion of fresh-cut pineapple was aseptically transferred from each tray to a sterile Stomacher bag (Seward Medical, London, UK) , diluted with 225 ml of Ringer buffer solution (Lab M Limited, Lanchashire, UK) and homogenized for 60 sec in a stomacher device (Lab Blender 400, Seward Medical). After the preparation of appropriate serial dilutions with Ringer solution, the following microbial determinations were performed: total mesophilic microbial

populations (total viable count, TVC) by the spread method on tryptic glycose yeast agar (Plate Count Agar, Biolife, Milan, Italy) , after incubation of plates at 25ºC for 72 h; yeast and moulds by the spread method on Rose Bengal Chloramphenicol agar (RBC, Lab M Limited) and incubation at 25ºC for 3-5 days; *Pseudomonas* spp. by spread method on pseudomonas agar base with selective supplement cephalothin-fucidin-cetrimide (CFC, Lab M Limited) and incubation at 25ºC for 48 h; lactic acid bacteria by pour method (with overlay) on de Man, Rogosa and Sharpe agar (MRS, Biolife) and incubation at 30ºC for 72 h; and bacteria of the Enterobacteriaceae family by pour method (with overlay) on violet red bile glucose agar (VRBG, Biolife) and incubation at 37ºC for 24 h. The results were expressed as the average (± standard deviation, *n*=8) log colony forming units per gram (log CFU/g) of fruit. The pH values of fruit samples were measured after the microbiological analysis, using a digital pH meter (RL150, Russell pH Cork, Ireland) with a glass electrode (Metrohm AG, Herisau, Switzerland).

4.2.3. Sensory analysis

Two staff members evaluated in duplicate the freshness rate of three different sensory features of the samples: odour, colour, and texture. For each sensory parameter, a score was given; 1 for fresh, 2 for intermediate, and 3 for unacceptable. Finally, the samples were classified in two classes: Class 1 for fresh (or acceptable) and Class 2 for non-acceptable pineapple samples. The intermediate samples were also classified in Class 2 to simplify the pipeline process and interpretation, since the samples with score 2 and 3 were commercially unacceptable.

4.2.4. FTIR spectroscopy

In parallel to microbiological analysis, FTIR spectra were collected using a ZnSe 45º HATR (Horizontal Attenuated Total Reflectance) crystal (PIKE Technologies, Madison, Wisconsin, United States) on a FTIR-6200 JASCO spectrometer (Jasco Corp., Tokyo, Japan) equipped with a triglycine-sulphate (TGS) detector and a Ge/KBr beamspliter. The samples were cut in small slices of such dimensions in order to cover the crystal and then, they were covered with a piece of aluminum foil. The spectral data were collected over the range of 4000–400cm^{-1} at room temperature (22 ± 2ºC), using the Spectra Manager™ Code of Federal Regulations (CFR) software version 2 (Jasco Corp.). Reference background spectra were collected every four

sample measurements by placing and acquiring a spectral measurement of the cleaned blank crystal. For both background and sample readings, 100 scans were accumulated at a nominal resolution of 4 cm^{-1}. The collection time for each sample spectrum was 2 min. At the end of each sample measurement, the crystal surface was cleaned with detergent, washed with distilled water, dried with lint-free tissue, cleaned with acetone and finally dried with lint-free tissue.

4.2.5. NIR spectroscopy

Spectra were acquired using a NIR spectrometer (SGS1900), developed by Fraunhofer IMPS (Institute für Photonische Mikrosysteme, Dresden, Germany). The NIR spectrometer covers a wavelength range from 1000 to 1900 nm, and it is comprised of a MEMS-based scanning grating for spectral dispersion and an uncooled InGaAs diode for detection (Pügner, Knobbe, Grüger & Schenk, 2012). The software "Quickstep" (Hiperscan GmbH, Dresden, Germany) was the operation software for the measurements. The organology and data acquisition procedures have been described in detail by Tsakanikas et al. (2018).

4.2.6. VIS and FLUO spectroscopy

The UV-VIS spectrometer used was the Hamamatsu C12880MA (Hamamatsu Photonics K.K., Shizuoka, Japan). The device has a spectral range from 850 to 340 nm and spectral resolution of 15 nm. It can be employed either for visible or fluorescence range spectroscopy by switching the mode of spectrometer and changing the settings. Specifically, the scan count was set at 10 and 3, while integration time at 250 μs and 100.000 μs for acquisition in visible and fluorescence mode, respectively. A UV filter with 400 nm cutoff wavelength was placed in the front of the spectrometer to obtain spectra only to the visible region. Before the samples measurements, a dark calibration and a reference acquisition are performed. Dark calibration is performed with the light source off placing the spectrometer on a dark surface for both modes. The white reference was performed with the light on using a white material (in our case a folded piece of paper) for visible mode and a non-fluorescent reflective reference standard (a black plastic surface) for fluorescence mode, respectively. The samples were placed in a petri dish covering the entire surface and 10 measurements (absorbance values)

were performed in different spots of each sample. The spectral values are expressed as the average of the 10 measurements for each wavelength after a normalization step.

4.2.7. Multispectral Image analysis

Multispectral images were captured using the VideometerLab device in 18 different wavelengths ranging from UV (405 nm) to short wave NIR (970 nm) (Carstensen & Hansen, 2003). The device has been commercialized by Videometer A/S. The spectral radiation is not continuous, but operates at wavelengths 405, 435, 450, 470, 505, 525,570, 590, 630, 645, 660, 700, 850, 870, 890, 910, 940 and 970 nm. The system is first calibrated radiometrically and geometrically using well-defined standard targets and a light setup is loaded based on the type of the product in each fresh form. The pineapple samples were placed in a Petri dish so that its surface was fully covered, and the dish was placed inside an Ulbricht sphere. The image acquisition and pre-treatment processes have been described previously in detail (Panagou et al., 2014) and have been implemented using the VideometerLab system software (version 2.12.39). For each image, the mean reflectance spectra (along with the standard deviation values) were calculated by averaging the intensity of pixels at each wavelength.

4.2.8. Data analysis

4.2.8.1. The Unscrambler software

Multivariate data analysis was initially carried out using the data analytics software, The Unscrambler© ver. 9.7 (CAMO Software AS, Oslo, Norway). Partial least squares regression (PLSR) was performed for the correlation between spectral data and microbial counts where, the spectral data were used as independent variables (X) and the TVC as dependent variables (Y). This method is considered suitable for spectroscopy datasets where the dimensionality problem exists (many variables but few samples) and also when the data show strong collinearity and noise (Wold et al., 2001; Mehmood et al., 2012; Gromski et al., 2015). For sensory features, since they are categorical variables, partial least squares discriminant analysis (PLSDA) was performed (Barker & Rayens, 2003).

For MSI, both the mean reflectance values and their standard deviations (in total 36 features) were used for model development, as it is considered that the second ones contain relevant and important information. For FTIR, PLSR was performed over the range 1800 and

870 cm^{-1}, since this range reflects the metabolic activity of microbial spoilage (Al-Jowder et al., 1999; Ellis et al., 2002; Di Egidio et al., 2009). For both FLUO and VIS spectral data, the range 700-400 nm was used for further analysis, while all the acquired spectra were used for model development in the case of NIR.

Prior to PLSR/PLSDA, several pre-processing techniques were tested on each dataset with the aim to minimize any irrelevant information such as noise, particle size deviations, scattering and drifting effects (Wang et al., 2015; Dixit et al., 2017; Suhandy & Yulia, 2017; Li et al., 2018). Specifically, for TVC prediction, standard normal variate (SNV) was the best pre-treatment for FTIR and FLUO, while VIS spectra were subjected to first derivative normalization with a second-order polynomial and a 9-point window. In the case of odour, MSI and NIR data were normalized using SNV, while for texture, only FLUO data were subjected to first derivative normalization with a second-order polynomial and a 9-point window. Additionally, the data from all sensors were mean-centered and scaled (1/SDEV).

The data derived from isothermal storage temperatures were used for the calibration process (training set) and those derived from dynamic temperature conditions for external validation (test set). During the calibration process, leave-one-out cross validation in parallel with Marten's uncertainty test was employed in order to eliminate the risk of over-fitting and test the predictive significance of the model, but also to select the significant X-variables (Westad & Martens, 2000; Wold et al., 2001). The significant independent variables were finally used for the construction of FTIR and FLUO models for texture assessment. In the case of NIR, a random data partitioning scheme (training set: 70%, test set: 30%) and the selection of significant wavelengths based on RFs regression ensemble (Breiman, 2001) were also used for TVC model development. Concerning RFs feature selection, spectra were initially normalized under the robust version of SNV, namely RNV (Barnes et al., 1989; Guo et al., 1999).

The prediction performance of the developed PLSR models for each type of spectral data was evaluated based on the following statistical parameters: slope (a), offset (b), correlation coefficient (r), the root mean square error (RMSE), the normalised RMSE (NRMSE) (Eq.1) and the coefficient of determination (R^2) of the linear regression between the predicted and measured microbial counts. For PLSDA models, the parameters for performance

evaluation were accuracy (Eq. 2), sensitivity (Eq. 3) and specificity (Eq. 4), where positive samples are the fresh or acceptable samples and negative samples are the non-acceptable ones.

$$\text{NRMSE} = \frac{\text{RMSE}}{\max(\text{DV}) - \min(\text{DV})}$$

(Eq.1)

where DV is the dependent variable (i.e. TVC).

$$\text{Accuracy} = \frac{\text{samples correctly predicted}}{\text{total number of samples}} \text{ x } 100$$

(Eq. 2)

$$\text{Sensitivity} = \frac{\text{true positive samples}}{\text{true positive samples} + \text{false negative samples}} \text{ x } 100$$

(Eq. 3)

$$\text{Specificity} = \frac{\text{true negative samples}}{\text{true negative samples} + \text{false positive samples}} \text{ x } 100$$

(Eq. 4)

4.2.8.2. SorfML platform

SorfML is a machine learning classification and regression analysis ranking system (www.SorfML.com). Specifically, it is a free web-platform able to automate the procedure of identifying the best machine learning method for comparing data from several analytical techniques and predict the freshness profiles as well as counts of microorganisms responsible for food spoilage. Using SorfML, users are able to securely upload raw experimental data collected using rapid and/or non-invasive analytical technologies (e.g. Multi/Hyper-spectral imaging, Electronic nose, Gas chromatography-Mass spectrometry) in CSV format, and apply various machine learning classification and regression modelling algorithms (e.g. SVM, Neural Network, RFs) in order to identify the best combination of analytical technology and

machine learning algorithm to predict given bacterial species or quality indices. An indicative workflow of the pineapple analysis followed in SorfML is presented in Figure 4.1.

The algorithms used in this study are kNN (Silverman & Jones, 1989), Ranger, a fast version of RFs (Wright & Ziegler, 2017), linear support vector machine (SVM Linear), radial support vector machine (SVM Radial) (Boser et al., 1992), PLSR (Wold et al., 2001) and PLSDA (Barker & Rayens, 2003). In order to generate each model, the dataset was randomly segmented into a training dataset for optimisation, with the 65% of the total samples, and a testing dataset for model validation with the left 35%. The different classes of data were equally represented in the training and test datasets. The spectral ranges used for each sensor were mentioned above (subsection 4.2.8.1), while the raw data were also centered and scaled, but no other data pre-processing was performed. It should be also noted that the NIR data were not included in the analyses performed with the SorfML platform.

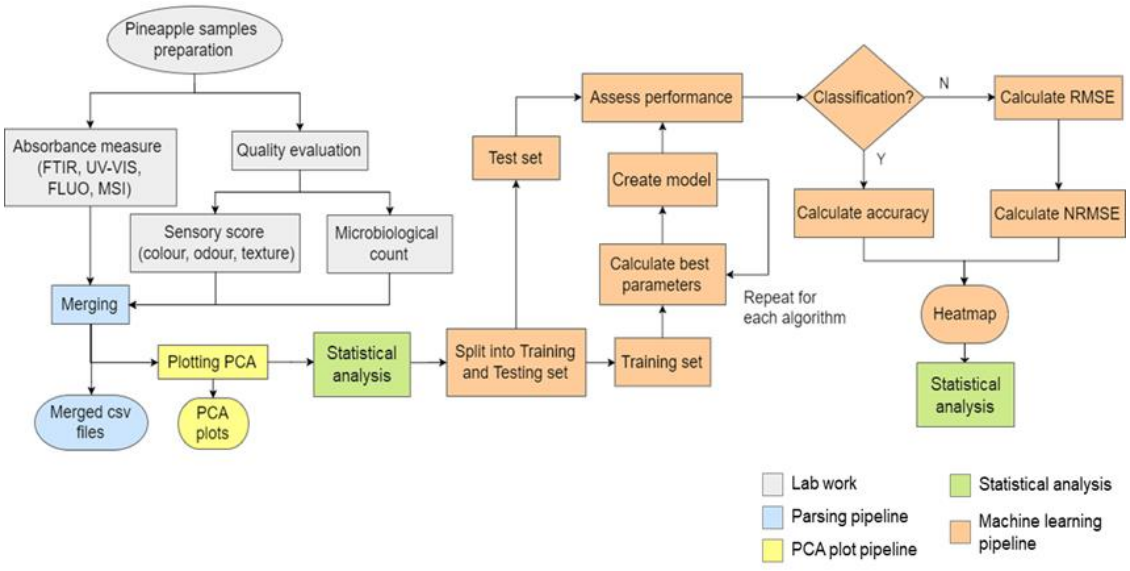


Figure 4.1. Workflow of SorfML. The workflow is divided into five sections with different colours corresponding to distinct parts of the applied methodology.

The predictive power of the developed models for sensory features was evaluated based on the accuracy parameter, but sensitivity and specificity parameters' values also are provided. For TVC models, the RMSE and NRMSE (Eq. 1) parameters for each analytical technology

were ranked in heatmaps, while R^2 values also were presented. Training performance was assessed using *k*-repeated fold cross-validation with a *k* value of 4 and a total number of 10 repetitions to choose the best parameters for each model.

4.3. Results

4.3.1. Microbiological spoilage and pH data

The initial level of TVC (mean ± standard deviation, n=8) was 5.09 ± 0.60 log CFU/g, while the final populations were 7.14 ± 0.50, 7.69 ± 0.40, 7.52 ± 0.30 and 7.91 ±0.20 log CFU/g during storage at 4, 8, 12 °C and under dynamic conditions, respectively. Similarly, the initial population of yeasts (mean ± standard deviation, n=8) was 5.01 ± 0.60 log CFU/g and reached at the end of storage 7.07 ± 0.50, 7.67 ± 0.40, 7.53 ± 0.30 and 7.78 ±0.30 log CFU/g at 4, 8, 12 °C and under dynamic conditions, respectively. The microbiological analysis indicated that the dominant spoilage group in pineapple was yeasts, with their counts coinciding with the total mesophiles' counts throughout storage (Figure 4.2). Bacterial population, for both Pseudomonadaceae and Enterobacteriaceae families, was very low with no more than 3 log CFU/g throughout storage at all the studied temperatures (data not shown). Moreover, the population of lactic acid bacteria was not detectable with common microbiological analyses, due to the overgrowth of yeasts on MRS agar plates even when cycloheximide was incorporated in the medium's formulation.

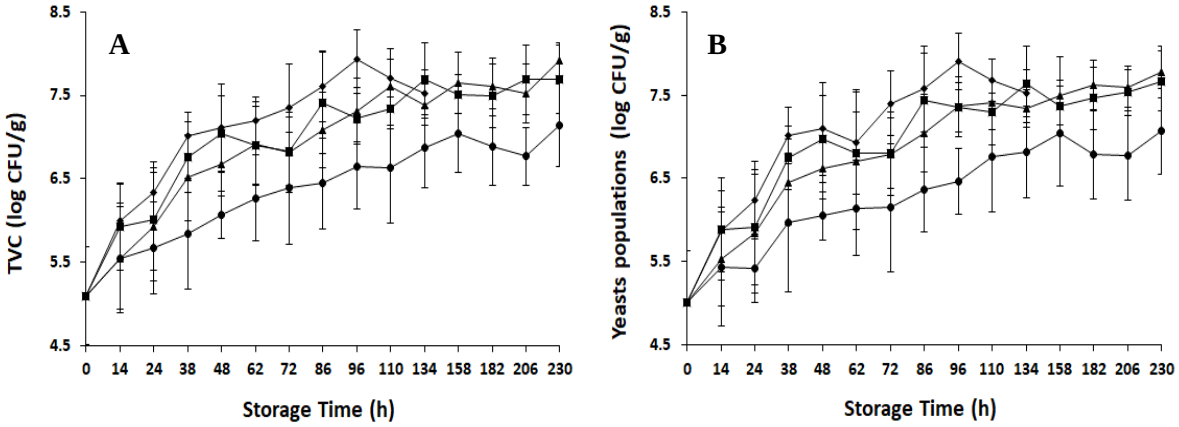


Figure 4.2. Total viable count (TVC) (A) and yeasts populations (B) in ready-to-eat pineapple during storage at 4 ºC (●), 8 ºC (■), 12 ºC (◆) and dynamic temperature conditions (▲). The

populations are expressed as means ± standard deviations (n=8 for FTIR, UV/VIS and MSI and n=4 for NIR).

As expected, the growth of the TVC and yeasts population was faster with increasing storage temperatures, while the growth monitored during storage at the dynamic temperature conditions resembled the one recorded during isothermal storage at 8 ^{o}C. As far as the pH is concerned, the values for pineapple were quite low and similar during storage at all studied temperatures (Table 4.1).

Table 4.1. The initial and final pH values of ready-to-eat pineapple during storage at different temperatures.

Storage temperature	**Initial pH**	**Final pH**
4 ^{o}C	3.45 ± 0.05	3.57 ± 0.04
8 ^{o}C		3.55 ± 0.07
12 ^{o}C		3.55 ± 0.17
Dynamic		3.52 ± 0.17

4.3.2 TVC models

Concerning TVC, the linear regression between the predicted (estimated) and the measured (observed) TVC values is presented in Table 4.2 and Figure 4.3 for The Unscrambler, and in Figure 4.4 for SorfML. The solid lines in Figures 4.3 and 4.4 are the ideal y=x lines, while the dashed ones determine the ±1 log unit area. Additionally, SorfML platform provides a ranking of performance (with RMSE and NRMSE values) of the various models developed for each analytical technology and machine learning algorithm, which is illustrated in a heatmap plot (Figure 4.5). In Table 4.4, the R^2 values are also presented for each model developed on SorfML software.

Starting with The Unscrambler, the presented NIR model (using random data partitioning and feature selection based on RFs) was the one with the best performance among the developed NIR models, but the poorest performance among the tested sensors. The corresponding results for NIR were not illustrated due to poor performance. Contrarily, the *a*

and *b* values of the linear regression between the estimated and the observed TVC values for FTIR, MSI, FLUO and VIS had a narrow range from 0.57 to 0.61 and 2.61 to 3.00, respectively.

Table 4.2. Performance metrics of the partial least squares reggresion models, based on the different analytical technologies (sensors), for the total viable count prediction of ready-to-eat pineapple on The Unscrambler software. (*a*: slope, *b*: offset, r: correlation coefficient, RMSE: root mean square error (log CFU/g), NRMSE: normalised root mean square error (%), R^2: coefficient of determination)

Data	***a***	***b***	**r**	**RMSE**	**NRMSE**	**R^2**
FTIR	0.61	2.61	0.70	0.59	13.43	0.43
NIR	0.27	5.05	0.44	0.52	11.84	0.15
MSI	0.61	2.78	0.74	0.53	12.06	0.54
FLUO	0.59	2.96	0.77	0.51	11.61	0.58
VIS	0.57	3.00	0.77	0.51	11.61	0.58

Additionally, the RMSE and NRMSE values were quite low for these sensors. Actually, the RMSE values were close to 0.5 log CFU/g. The R^2 values were under 0.6 for the four sensors, while the r value ranged from 0.70 to 0.77. The FLUO and VIS models exhibited the highest r and R^2 values, as well as the lowest RMSE and NRMSE values. The number of components selected for each PLSR model are presented in Table 4.3.

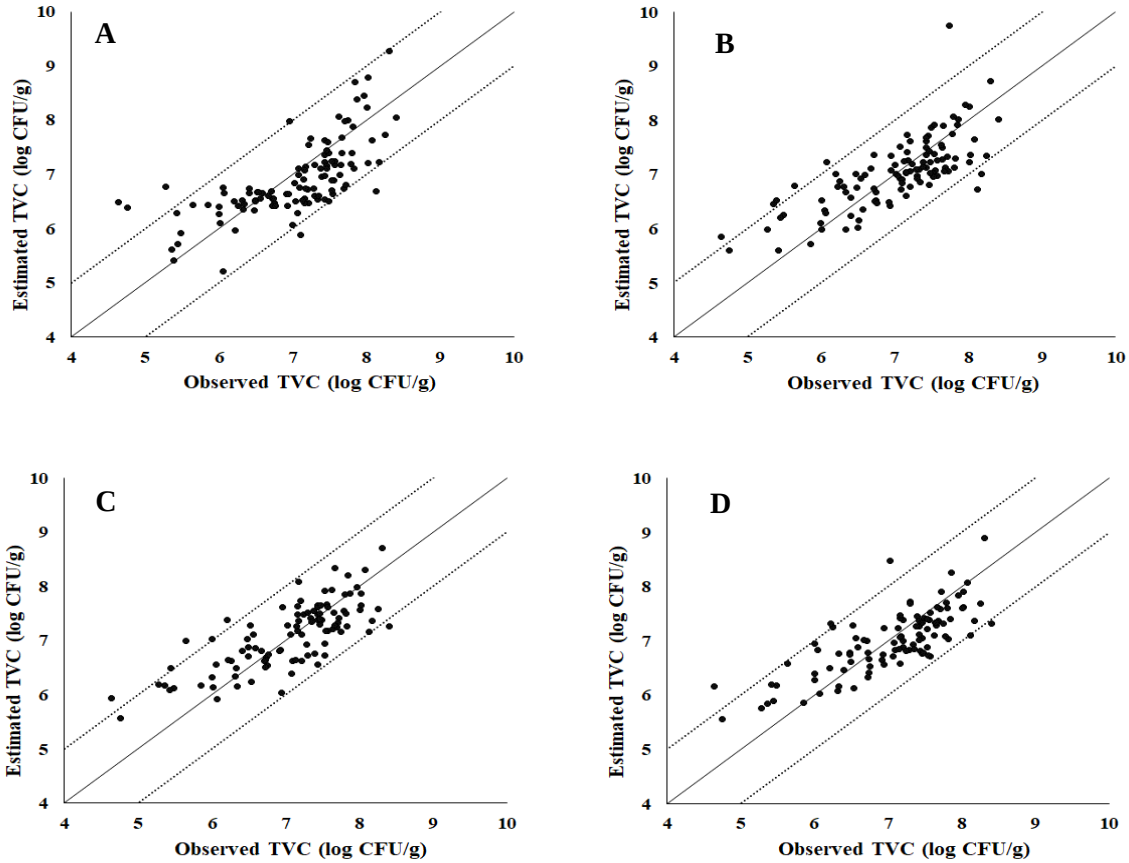


Figure 4.3. Linear regression between the predicted and the measured by the partial least squares regression model, total viable count (TVC) values of prediction based on FTIR (A), MSI (B), FLUO (C) and VIS (D) data for ready-to-eat pineapple, using The Unscrambler software (solid line: the ideal y=x line; dashed lines: the ± 1 log unit area). The corresponding results for NIR were not illustrated due to poor performance.

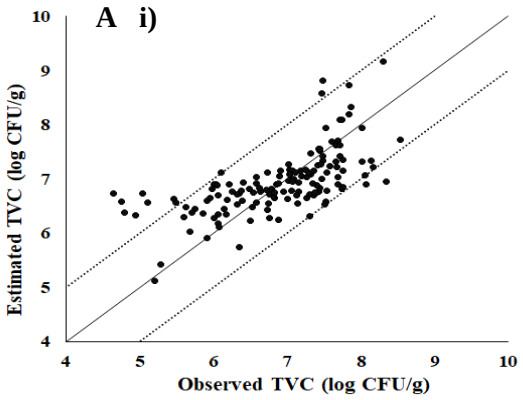
In SorfML, RMSE values for all the studied sensors and algorithms were also below 1 log CFU/g. For almost all sensors, PLSR and SVM Linear algorithms exhibited the lowest values of RMSE (and NRMSE), with a range from 0.58 to 0.64 log CFU/g and the highest R^2 values, with a range from 0.41 to 0.50. On the other hand, Ranger, kNN and SVM Radial exhibited the highest RMSE values, above 0.65 log CFU/gr, as well as the lowest R^2 values. The regression lines between the estimated and the measured TVC values for non-linear models are not presented in Figure 4.4. The best model accuracy was achieved through the PLSR model combined with FLUO spectral data, showing the lowest RMSE/NRMSE values

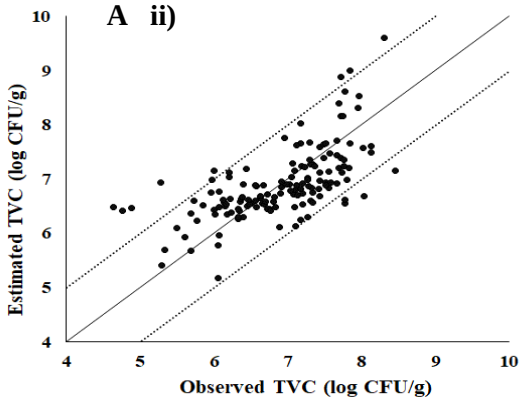
and the highest R^2 values. The tuning parameters selected for each model in SorfML software are presented in Table 4.5.

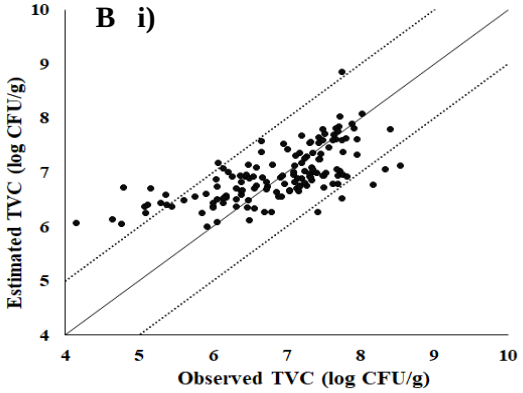
Table 4.3. Number of latent variables for partial least squares regression (PLSR) and partial least squares discriminant analysis (PLSDA) models developed for the total viable count and sensory features prediction of ready-to-eat pineapple on The Unscrambler software.

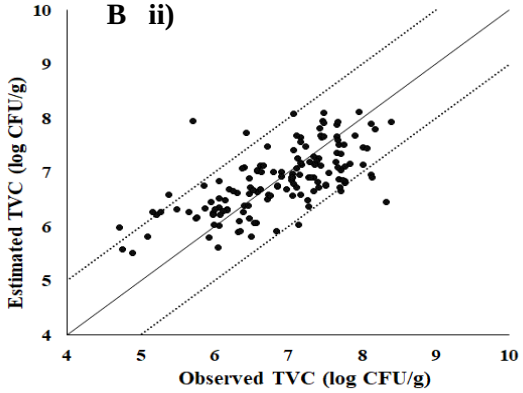
Feature	Sensor	PLSR/PLSDA nc	Feature	Sensor	PLSR/PLSDA nc
TVC	FTIR	7	**Odour**	FTIR	6
	NIR	3		NIR	1
	MSI	11		MSI	10
	FLUO	7		FLUO	6
	VIS	12		VIS	7
Colour	FTIR	7	**Texture**	FTIR	5
	MSI	5		MSI	1
	FLUO	6		FLUO	3
	VIS	7		VIS	6

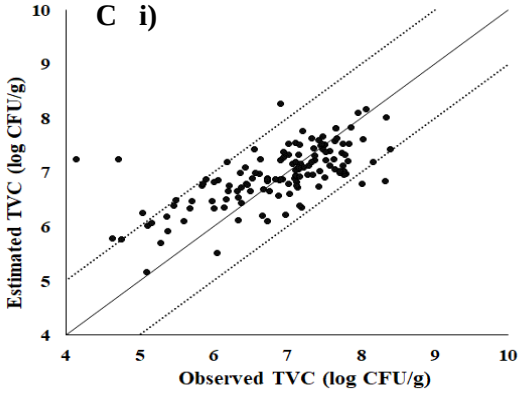
*nc: number of components

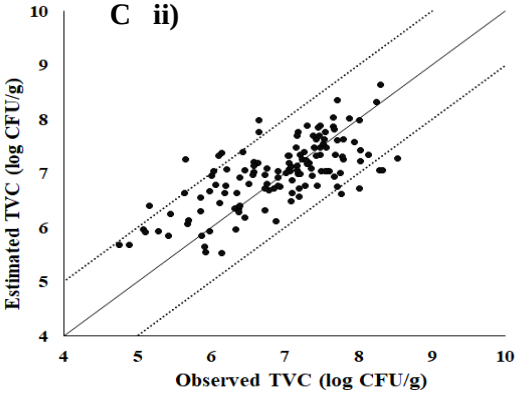

A i)
Estimated TVC (log CFU/g)
Observed TVC (log CFU/g)
4
5
6
7
8
9
10


A ii)
Estimated TVC (log CFU/g)
Observed TVC (log CFU/g)
4
5
6
7
8
9
10


B i)
Estimated TVC (log CFU/g)
Observed TVC (log CFU/g)
4
5
6
7
8
9
10


B ii)
Estimated TVC (log CFU/g)
Observed TVC (log CFU/g)
4
5
6
7
8
9
10


C i)
Estimated TVC (log CFU/g)
Observed TVC (log CFU/g)
4
5
6
7
8
9
10


C ii)
Estimated TVC (log CFU/g)
Observed TVC (log CFU/g)
4
5
6
7
8
9
10

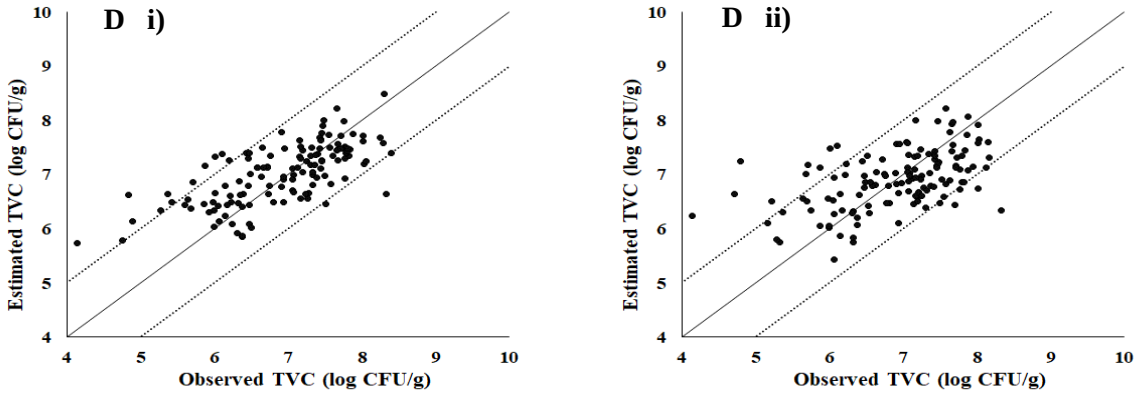


Figure 4.4. The linear regression between the predicted and the measured total viable count values by i) support vector machine (SVM) Linear and ii) partial least squares regression (PLSR) models, based on FTIR (A), MSI (B), FLUO (C) and VIS (D) data for ready-to-eat pineapple, using SorfML software (solid line: the ideal y=x line; dashed lines: the ± 1 log unit area).

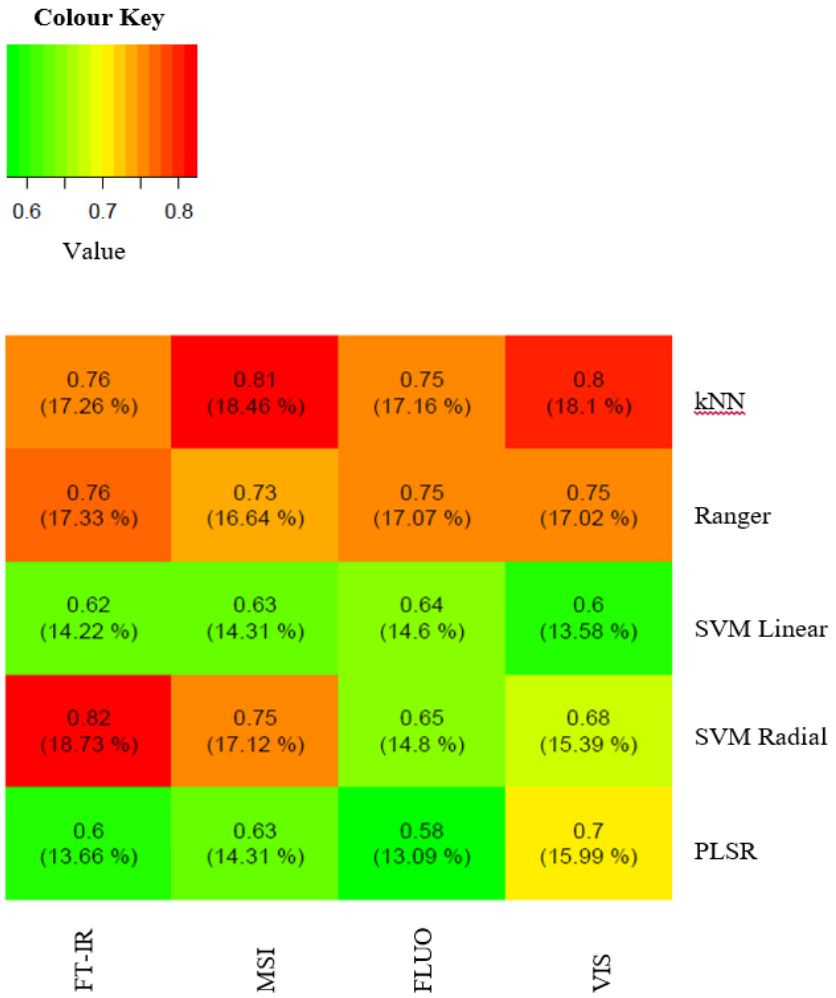


Figure 4.5. Performance heatmap of the different models developed for each analytical technology for the total viable count prediction of ready-to-eat pineapple. In the heatmap, the rows correspond to the five different algorithms, while the columns to the four different analytical technologies. The root mean square error (RMSE: log CFU/g) and normalized root mean square error (NRMSE: %) values are presented. The colour key depicts the extreme intensity for the extreme values. The colour key begins from green (higher performance) to red (lower performance).

Table 4.4. The coefficient of determination (R^2) of the models derived from the various combinations of the different analytical technologies (sensors) and algorithms, for the total viable count prediction of ready-to-eat pineapple using the SorfML software.

Sensor	**Algorithm**	**R^2**	**Sensor**	**Algorithm**	**R^2**
	kNN	0.07		kNN	0.22
	Ranger	0.16		Ranger	0.19
FTIR	SVM Linear	0.42	FLUO	SVM Linear	0.46
	SVM Radial	0.12		SVM Radial	0.39
	PLSR	0.42		PLSR	0.5
	kNN	0.09		kNN	0.08
	Ranger	0.13		Ranger	0.21
MSI	SVM Linear	0.44	VIS	SVM Linear	0.48
	SVM Radial	0.12		SVM Radial	0.26
	PLSR	0.41		PLSR	0.27

Table 4.5. Tuning parameters used in the development of the different models for the total viable count and sensory features prediction of ready-to-eat pineapple using the SorfML software.

Model	Sensory Feature	kNN	Ranger			SVM Linear	SVM Radial		PLSR/ PLSDA
		k	mtry	splitrule	min.node .size	c	sigma	c	nt
FTIR	TVC	13	963	variance	5	1	0.03	4	7
	Colour	5	7	extratrees	1	1	0.02	1	8
	Odour	23	620	gini	1	1	0.02	128	7
	Texture	9	11	extratrees	1	1	0.01	0.25	7
MSI	TVC	17	31	variance	5	1	0.05	2	11
	Colour	23	2	extratrees	1	1	0.05	0.25	5
	Odour	11	26	gini	1	1	0.06	16	14
	Texture	15	2	extratrees	1	1	0.06	0.5	14
FLUO	TVC	33	122	extratrees	5	1	0.01	8	7
	Colour	9	67	gini	1	1	0.01	0.5	7
	Odour	17	57	gini	1	1	0.01	4	5
	Texture	11	113	gini	1	1	0.01	32	11
VIS	TVC	27	94	variance	5	1	0.02	4	6
	Colour	5	2	extratrees	1	1	0.02	0.5	5
	Odour	5	104	gini	1	1	0.02	16	11
	Texture	5	85	gini	1	1	0.02	4	15

kNN*: k=number of neighbours considered,*

Ranger: mtry=number of variables to possibly split at in each node, splitrule=splitting rule, min.node.size=minimal node size,

SVM: c=cost of constraints violation, sigma=scale parameter of the hypothesized (zero-mean) Laplace distribution estimated by maximum likelihood

PLSR/PLSDA: nt=number of components

4.3.3. Sensory models

Due to the limited data used for sensory analysis, the corresponding results are presented as supplementary information. Table 4.6 presents the number of acceptable and non-

acceptable (in terms of freshness) samples for the three sensory features. It should be noted that for colour and texture, the number of fresh samples are four and three times higher than that of non-acceptable, while for odour, the samples are more equally distributed to both classes. Therefore, it seems that odour is the quality attribute that better reflects the spoilage of pineapple compared to the colour and texture. The prediction performances of the models developed for each one of the tested analytical technologies are summarized in Table 4.7 for The Unscrambler and the number of components selected for each PLSR model are presented in Table 4.3. The accuracy ranking of models generated for sensory features using the SorfML software is illustrated in Figure 4.6, while Table 4.8 provides the corresponding values of sensitivity and specificity.

For The Unscrambler, none of the models reached 90% of accuracy. The models for texture prediction had the lowest accuracy values, while for colour and odour, exceeded 80% for almost all the sensors (apart from FTIR and NIR in the case of odour).

Table 4.6. Total number of the measured fresh (acceptable) and non-acceptable samples for each sensory feature.

Sensor	Sensory Feature	Number of Fresh	Number of Non-Acceptable	Total Number
FTIR / MSI	Colour	333	91	424
	Odour	147	277	
	Texture	314	110	
FLUO / VIS	Colour	314	78	392
	Odour	143	249	
	Texture	292	100	

However, in the case of colour, the satisfactory accuracy values of models were not followed by similar sensitivity and specificity values. As far as the odour prediction is concerned, FLUO model is the only model, which has the highest values of both sensitivity (88.46%) and

specificity (85.90%), and the highest value of accuracy (86.54%) among the analytical technologies.

Table 4.7. Performance metrics of the partial least squares discriminant analysis models based on the different analytical technologies (sensors), for the sensory features prediction of ready-to-eat pineapple using The Unscrambler software.

Sensor	Sensory Feature	Accuracy (%)	Sensitivity (%)	Specificity (%)
FTIR	Colour	83.04	97.62	39.29
	Odour	76.79	50.00	84.88
	Texture	72.32	90.00	28.13
NIR	Odour	71.43	18.18	90.32
MSI	Colour	85.71	100	42.86
	Odour	83.04	61.54	89.53
	Texture	72.32	95.00	15.63
FLUO	Colour	84.62	93.67	56.00
	Odour	86.54	88.46	85.90
	Texture	70.19	90.54	20.00
VIS	Colour	81.73	98.73	28.00
	Odour	83.65	57.69	92.31
	Texture	68.27	90.54	13.33

In SorfML, the accuracy percentage of all the models generated for all of the three features, did not exceed 89.23%. As far as the colour and texture are concerned, the models with accuracy values above 80% showed also high sensitivity and low specificity. For odour, FLUO combined with PLSDA, SVM Radial and SVM Linear, but also MSI combined with SVM Linear, showed the highest accuracy values. The combinations with the best accuracy, sensitivity and specificity values were FLUO and PLSDA, as well as MSI and SVM Linear.

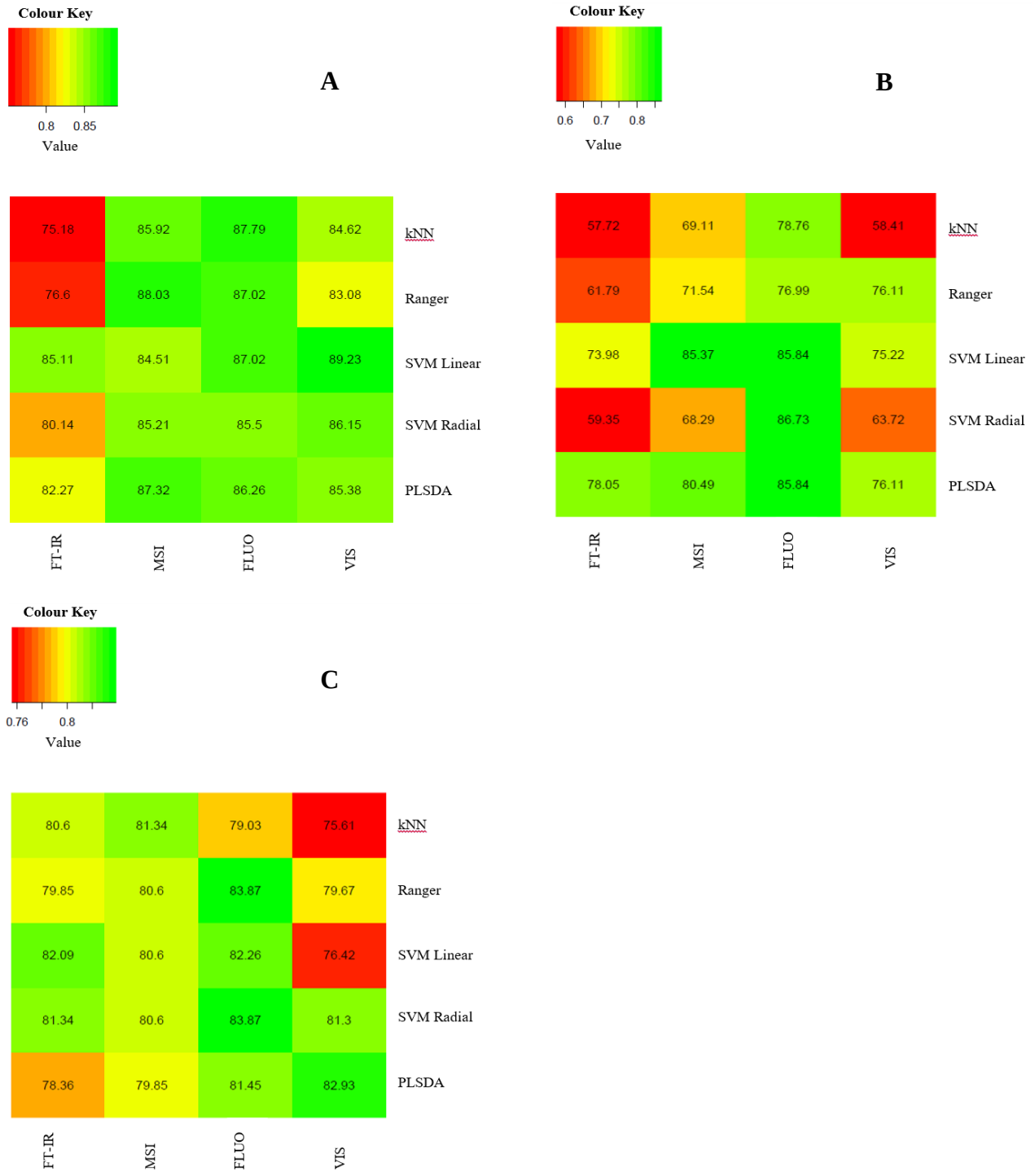


Figure 4.6. Performance heatmaps of the different models developed for each analytical technology for the Colour (A), Odour (B) and Texture (C) prediction of ready-to-eat pineapple. In each heatmap, the rows correspond to the five different algorithms, while the columns to the four different analytical technologies. The accuracy (%) values are presented for the

sensory features. The colour key depicts the extreme intensity for the extreme values. The colour key begins from red (lower performance) to green (higher performance).

4.4. Discussion

The introduction of spectroscopic and optical sensing (computer vision) methods in food science and their growing application in a wide range of food products is becoming a clear trend during the last decade. Although these analytical techniques offer rapid and in many cases efficient answers in a non-destructive way, the management and analysis of data and the interpretation of results are still a great challenge (Zhou et al., 2019).

Data analysis tools is a set of technology that enable users to analyze and visualize data in order to identify trends and correlations with the goal of supporting decision making. Therefore, data analysis software is considered to be a central requirement for any sector / business (Vassakis et al., 2018). Although a large number of tools are available nowadays, only a limited number of these have proven to be useful for food science. The two applied statistical software, The Unscrambler and SorfML, provide a user-friendly and easy-to-use interface for analyzing large experimental datasets. These automated software also provide the opportunity for scientists, let alone food microbiologists with limited statistical and mathematical knowledge, to perform the challenging task of data mining and predictive modelling. However, the users are responsible for learning the advantages and the limitations of each tool and also realize that conflicting outputs are even possible (Nunes et al., 2015).

The Unscrambler software provides the option of limited in number and only linear algorithms, namely PLSR and PLSDA. On the other hand, SorfML platform offers the wide option of choosing linear and non-linear algorithms often used in data analysis. The approach of using different algorithms for the same data, allows comparing the performance of each created model and evaluate their suitability for different scenarios (Estelles-Lopez et al., 2017). Apart from algorithms, the big differences between the two data analytics tools lie in the partitioning of data in training and external validation sets, the pre-processing of data and the cross-validation method.

Starting with data partitioning when using The Unscrambler, model testing was performed on samples stored in dynamic temperature conditions, since these data include the information from all temperatures and are also considered as a simulation of real life in the

food supply chain (Tsakanikas et al., 2018). On the other hand, in SorfML, the data were randomly segmented ensuring that all temperature and storage time groups were equally represented in both datasets and finally provided a less biased selection of sample. Furthemore, using The Unscrambler, various pre-processing methods were tested and those with the best results were finally applied. The purpose was to remove the irrelevant information from data and facilitate their interpretation. However, in SorfML, the data were subjected to the minimum treatment, with the aim not to lose important information. As far as the cross validation was concerned, leave-one-out (or full-cross) was used in the The Unscrambler, while *k*-fold cross-validation was used in SorfML. Full-cross validation is a common method, where all samples are used in an exhaustive way providing repeatability of the results, but also may lead to over-optimistic results. On the other hand, a *k*-fold partitioning could potentially result in folds where samples are not represented in an equal manner (Ropodi et al., 2016).

Besides all these different approaches, the summarized results indicated similar trends about the ability of the studied sensors and algorithms to assess the quality of RTE pineapple. Specifically, for TVC assessment, all acquired spectral data (apart from NIR in The Unscrambler and VIS in SorfML) combined with PLSR algorithm showed satisfactory performance in both The Unscrambler and SorfML. In The Unscrambler, the best models with slight differences compared to the others, were the ones based on FLUO and VIS data. In SorfML, FLUO and PLSR combination also exhibited the best performance with slight differences compared to the other PLSR models. Apart from the PLSR algorithm, the SVM Linear combined with every sensor was also appropriate for TVC prediction. Contrarily, non-linear algorithms, tested in SorfML, did not manage to predict the spoilage of RTE pineapple. In the case of NIR, the model performance was poor compared to the other sensors, indicating that this analytical technology is not appropriate for the evaluation of the microbiological spoilage of pineapple.

An important observation on these results, is that the R^2 values were quite low even for the best model performances. It could be argued that according to these low R^2 values the prediction performance is poor. However, the variability present in the data strongly influences this parameter. Indeed, the variability among the experimental replications (batches), as well as between the two biological replicates (duplicate samples) of the same experimental replication was rather high (data not shown). This excessive variability is very

common in plant origin products due to their natural variability and the strong impact of various factors such as, cultivar, geographical region, agricultural practices and post-harvest factors (i.e. processing, storage) results in large spectral and microbiological variations (Rico et al., 2007; Zhang et al., 2018). Although, there is no guarantee that a high value of R^2 is indicative of the 'goodness of fit', a low value of this parameter is most problematic when precise predictions are necessary (Granato et al., 2014). In food microbiology, 0.5 log deviations in microbial counts are common even within the same laboratory, but also RMSE values under 1 log CFU/g are acceptable for food microbiology applications (Tsakanikas et al., 2018). In the present study, most of the predictions of the best-performing models presented in this study were within this limit. However, large datasets containing more variability should be obtained in order to develop more universal and accurate models for better assessment of the studied phenomenon.

For sensory features, the results were presented in order to reveal a potential trend which would be helpful for further investigation. The results derived from the two platforms showed similar conclusions. Regarding the low number of non-acceptable over the fresh samples for colour and texture, the models generated by both software were biased over the fresh samples. Consequently, the presence of a dataset with a more balanced proportion of the two classes, could possibly result in better model performances and safer conclusions for colour and texture assessment. For odour, the most representative of spoilage attribute, FLUO data and PLSDA algorithm emerged as one of the most appropriate combinations for both tools. Additionally, SorfML software indicated that the combination of MSI and SVM Linear may be also appropriate for odour assessment in pineapple.

Table 4.8. The specificity and sensitivity values of the partial least squares discriminant analysis models, derived from the various combinations of the different analytical technologies (sensors) and algorithms, for the sensory features prediction of ready-to-eat pineapple on SorfML software. The accuracy values are presented in heatmaps.

Sensor	Algorithm	Sensory Feature	Sensitivity (%)	Specificity (%)	Sensor	Algorithm	Sensory Feature	Sensitivity (%)	Specificity (%)
FTIR	kNN	Colour	91.15	10.71	MSI	kNN	Colour	95.61	46.43
		Odour	24.00	80.82			Odour	58.00	76.71
		Texture	99.06	10.71			Texture	97.17	21.43
	Ranger	Colour	92.04	14.29		Ranger	Colour	99.12	42.86
		Odour	46.00	72.60			Odour	48.00	87.67
		Texture	93.40	28.57			Texture	97.17	17.86
	SVM Linear	Colour	99.11	28.57		SVM Linear	Colour	100.00	21.43
		Odour	74.00	73.97			Odour	84.00	86.30
		Texture	99.06	17.86			Texture	93.40	32.14
	SVM Radial	Colour	97.35	10.71		SVM Radial	Colour	93.86	50.00
		Odour	36.00	75.34			Odour	50.00	80.82
		Texture	100.00	10.71			Texture	98.11	14.29
	PLSDA	Colour	92.92	39.29		PLSDA	Colour	100.00	35.71
		Odour	82.00	75.34			Odour	78.00	82.19
		Texture	96.23	10.71			Texture	91.51	35.71

Table 4.8 (continued)

Sensor	Algorithm	Sensory Feature	Sensitivity (%)	Specificity (%)	Sensor	Algorithm	Sensory Feature	Sensitivity (%)	Specificity (%)
FLUO	kNN	Colour	98.15	39.13	VIS	kNN	Colour	95.33	34.78
		Odour	83.67	75.00			Odour	57.14	59.38
		Texture	95.00	12.50			Texture	87.88	25.00
	Ranger	Colour	97.22	39.13		Ranger	Colour	96.26	21.74
		Odour	67.35	84.38			Odour	65.31	84.38
		Texture	97.00	29.17			Texture	92.93	25.00
	SVM Linear	Colour	100.00	26.09		SVM Linear	Colour	95.33	60.87
		Odour	79.59	90.63			Odour	73.47	76.56
		Texture	100.00	8.33			Texture	86.87	33.33
	SVM Radial	Colour	97.22	30.43		SVM Radial	Colour	97.20	34.78
		Odour	79.59	92.19			Odour	63.27	64.06
		Texture	97.00	29.17			Texture	97.98	12.50
	PLSDA	Colour	99.07	26.09		PLSDA	Colour	98.13	26.09
		Odour	85.71	85.94			Odour	81.63	71.88
		Texture	93.00	33.33			Texture	94.95	33.33

4.5. Conclusions

The implementation of different data analytics tools requires good knowledge of their range of applications and limitations and the results should always be evaluated critically. According to the the results of this study, both The Unscrambler and SorfML revealed similar trends for the various sensors. Specifically, the assessment of pineapple spoilage could be potentially achieved by the various spectroscopy-based technologies studied, namely FTIR, FLUO, VIS as well as MSI, but not the NIR sensor. As far as the sensory features are concerned, the odour could be possibly assessed by FLUO spectroscopy, provided that a more integrated and representative analysis is conducted in the future. The findings of this work also demonstrate that the comparative application and evaluation of various algorithms may lead to new options and more reliable results.

However, the design of large-scale experiments containing a satisfactory level of variability and the conduction of further research on data analysis including significant feature (wavelength/wavenumber) selection and data fusion strategies are needed. All the above will contribute to the development of even more robust and accurate models for microbiological spoilage of RTE pineapple aiming to the future replacement of the conventional microbiological methods.

Acknowledgments

This work has been supported by the project “PhasmaFOOD”, funded from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 732541.

Chapter 5

Application of spectroscopy-based technologies for evaluating microbiological spoilage of ready-to-eat leafy vegetables

Part of this chapter has been submitted in International Journal of Food Microbiology 'Spectroscopy and imaging technologies coupled with machine learning for the assessment of the microbiological spoilage associated to ready-to-eat leafy vegetables'. ***Evanthia Manthou****, Apostolos Karnavas, Lemonia-Christina Fengou, Anastasia Bakalaki, Alexandra Lianou, Panagiotis Tsakanikas, George –John E. Nychas*

Chapter 5

Abstract

Based on both new and previously collected experimental data (Tsakanikas et al., 2018), the present study provides a comparative assessment of sensors and machine learning approaches for evaluating the microbiological spoilage of ready-to-eat leafy vegetables (rocket and baby spinach) stored under passive modified atmosphere packaging (MAP). Fourier transform infrared (FTIR), near-infrared (NIR), visible (VIS) spectroscopy and multispectral imaging (MSI) were used. Two data partitioning approaches and two algorithms, namely partial least squares regression and support vector regression (SVR) based on radial basis function (RBF) kernel, were evaluated. Moreover, the abovementioned spectroscopy-based technologies were applied for microbiological spoilage assessment of baby spinach stored under active MAP. Concerning baby spinach, when model testing was performed on samples randomly selected, the performances were better or similar to the one attained when testing was performed based on dynamic temperatures data, depending on the applied analytical technology. The two applied algorithms yielded similar model performances for the majority of baby spinach cases. Regarding rocket, the random data partitioning approach performed considerably better results in almost all cases of sensor/algorithm combination. Furthermore, SVR algorithm resulted in considerably or slightly better model performances for the FTIR, VIS and NIR sensors, depending on the data partitioning approach. However, PLSR algorithm provided better models for the MSI sensor. Overall, the microbiological spoilage of baby spinach stored under passive MAP was better assessed by models derived mainly from the VIS sensor, while FTIR and MSI were more suitable in rocket. On the contrary, FTIR and MSI were more appropriate for baby spinach stored under active MAP. The results indicated that a distinct sensor and computational data analysis workflow is needed, not only for each vegetable type but also, for the different packaging conditions.

5.1. Introduction

In the last decades, fresh-cut produce supply is one of the growing industries of the agro-food sector, worldwide (Sandhya et al., 2010). Ready-to-eat (RTE) vegetable salads constitute a major category of the fresh-cut market, since their production and variety show a progressive increase over the years, due to important changes in consumers' lifestyle (Arienzo et al., 2020). Nowadays, modern food preferences include healthy, nutritious, minimally processed, but at the same time convenient food products (Qadri et al. 2015; Chaudry et al., 2018).

Despite the health benefits of consuming raw vegetables, these commodities are considered highly perishable (Leff & Fierer, 2013). The mild processing operations applied to RTE vegetable salads accelerate the physiological deterioration process, the evolution of biochemical changes, while at the same time enhancing microbial contamination and degradation of microbial etiology (Lunadei et al., 2012; Qadri et al., 2015). Since fresh produce commodities are naturally contaminated with a large amount of spoilage microorganisms, proliferation of spoilage microorganisms may considerably affect the already limited shelf-life of these commodities resulting in remarkably high economic losses (Granato et al., 2018; Giannoglou et al., 2020). Therefore, the development of effective and rapid food spoilage detection methods may be very supportive for the purpose of quality management in the fresh-cut produce industry (Wang et al., 2018).

The exploration for rapid analytical techniques has become an important research topic for the assessment of food safety and quality. Compared to conventional microbiological methods, spectroscopy and imaging approaches have become popular and rather attractive due to minimal sample preparation, non-destructive sampling, rapid data acquisition, as well as in-, on- and at-line detection potential (Nychas et al., 2016; Wang et al., 2018). However, the massive amount of data derived from such analytical approaches pose an important challenge for data mining and analysis purposes, usually demanding a multidisciplinary approach. For this reason, such technologies are commonly coupled with advanced computational methods including chemometrics, machine learning approaches and artificial intelligence techniques (Jollife & Cadima, 2016; Ropodi et al., 2016; Truong et al., 2019; Tsakanikas et al., 2020).

Advances in data science have introduced various machine learning approaches, applied in tandem with the above analytical technologies for the qualitative and/or quantitative

evaluation of the microbiological spoilage of various food products, including mainly meat products (Argyri et al., 2013; Grewal et al., 2015; Oto et al., 2013; Panagou et al., 2014; Fengou et al., 2019a; Spyrelli et al. 2020) but also fish (Duan et al., 2014; He & Sun, 2015a; He & Sun, 2015b; Saraiva et al., 2016; Fengou et al., 2019b), and to a lesser extent, dairy products (Nicolaou & Goodacre, 2008; Lianou et al., 2019), fruits (Di Egidio et al., 2009; De Sousa Marques et al., 2013; Al-Holy et al., 2015; Manthou et al., 2020) and vegetables (Wang et al., 2010; Tsakanikas et al., 2018). However, it is often challenging to select the appropriate combinations of analytical technologies and machine learning approaches and, thus, comparative evaluation for choosing the best approach for specific type of data is still needed (Cozzolino et al., 2015; Estelles-Lopez et al., 2017).

Beyond this, the application of analytical technologies for the evaluation of the microbiological quality of fresh-cut produce is limited in the scientific literature. In this context, the objective of the present study was the comparative assessment of non-invasive sensors and machine learning approaches for evaluating the microbiological spoilage of RTE leafy vegetables when stored at passive and/or active modified atmosphere packaging (MAP) conditions. For this purpose, four analytical technologies, based on Fourier transform infrared (FTIR), near-infrared (NIR) and visible (VIS) spectroscopies as well as multispectral imaging (MSI), were used for the assessment of microbiological quality of RTE rocket and baby spinach salads. Although part of the experimental data (derived from FTIR, NIR and VIS) was previously generated by Tsakanikas et al. (2018), different computational analysis was applied herein, while a considerable dataset enrichment via new experimental measurements also was performed. With reference to data analysis, two machine learning algorithms and two distinct data partitioning approaches were utilized. To the best of our knowledge, this is the first time that different analytical technologies and machine learning approaches are comparatively evaluated for the purpose of assessing the microbiological quality of vegetable products, with the ultimate goal of choosing the optimum combination in terms of model prediction performance.

5.2. Materials and Methods

5.2.1. Samples and storage conditions

Fresh-cut and washed RTE rocket and baby spinach salads were purchased from a local manufacturer (Athens) and transported to the laboratory within 24 h from production. The vegetables were packaged in plastic bags each one containing 85-g or 125-g portions of rocket or baby spinach leaves, respectively. Baby spinach and rocket samples were stored in their original commercial packaging under passive MAP due to their respiration, while some samples of baby spinach were also stored under active MAP. The latter were enclosed in plastic packages (length: 25 cm, width: 25 cm, thickness: 90 μm, permeability of ca. 25, 90, 6 cm^3 m^{-2}day^{-1}bar^{-1}(1 bar= 10^5Pa), at 20 °C and 50% RH for CO_2, O_2 and N_2, respectively) and flushed with 95%O_2 and 5%CO_2 using a HenkoVac 1900 Machine (Howden Food Equipment BV, The Netherlands). According to Allende et al. (2004), the super atmospheric O_2 treatment seems to contribute to lower tissue electrolyte leakage and higher overall quality scores and reduce aerobic mesophilic growth. The samples from both packaging conditions were stored at three different constant temperatures, namely 4, 8 and 12 ºC, in high precision (±0.5) programmable incubators (MIR-153, Sanyo Electric Co., Osaka, Japan). The baby spinach and rocket samples stored under passive MAP were also stored under dynamic temperature conditions (8 h at 4 ºC, 8 h at 8 °C and 8 h at 12 ºC), The temperature was recorded at 15-minutes intervals using electronic temperature devices (COX TRACER®, Cox Technologies Inc., Belmont, NC, USA). At regular time intervals, depending on the applied storage temperature and the packaging conditions, duplicate samples (originating from different packages) were subjected to the following analytical procedures: (i) microbiological analyses and pH measurements; (ii) FTIR spectroscopy measurements; (iii) VIS spectroscopy measurements, (iv) NIR spectroscopy measurements, and (v) MSI acquisition.

In the case of passive MAP, two independent experimental replications were conducted for baby spinach and three for rocket. A total of 265 rocket samples and 216 baby spinach samples were analysed. In the third experimental replication of rocket, dynamic conditions were not applied. Part of the above experimental data (referring to the FTIR, NIR and VIS sensors) were previously utilized in the development of a unified spectra analysis workflow (Tsakanikas et al., 2018). Nonetheless, the originally generated dataset has been enriched both qualitatively (via the addition of MSI) and quantitatively (the third experimental replication in

rocket), whereas a distinct computational analysis procedure has been embraced in the framework of the present study. With regard to baby spinach stored under active MAP, two independent experimental replications were conducted and the microbiological spoilage was assessed through FTIR and MSI; a total of 120 samples were analysed for both sensors, since no dynamic conditions were applied.

5.2.2. Microbiological analyses, pH and gas composition measurements

Prior to microbiological analyses, the O_2/CO_2 composition inside the packages was measured using a headspace gas analyzer (CheckMate 9900, PBI Dansensor, Denmark). Subsequently, a 25-g portion of fresh-cut vegetable was aseptically transferred from each plastic bag to a sterile Stomacher bag (Seward Medical, London, UK), diluted with 225 ml of quarter-strength Ringer's solution (Lab M Limited, Lanchashire, UK) and homogenized for 60 sec in a stomacher device (Lab Blender 400, Seward Medical). Appropriate serial decimal dilutions were prepared in Ringer's solution and were surface plated on tryptic glycose yeast agar (Plate Count Agar, Biolife, Milan, Italy) for the determination of the total mesophilic microbial population (total viable count, TVC). Presumptive *Pseudomonas* spp. were also enumerated by surface plating on pseudomonas agar base with selective supplement cephalothin-fucidin-cetrimide (CFC, Lab M Limited). The plates were incubated at 25 ^{o}C for 72 h and 48 h for TVC and *Pseudomonas* spp., respectively. Additional microbial determinations were performed, including lactic acid bacteria (LAB) by pour method (with overlay) on de Man, Rogosa and Sharpe agar (MRS, Biolife) and incubation at 30^{o}C for 72 h; and bacteria of the Enterobacteriaceae family by pour method (with overlay) on violet red bile glucose agar (VRBG, Biolife) and incubation at 37^{o}C for 24 h. The results were expressed as the average (± standard deviation) log colony forming units per gram (log CFU/g) of vegetable. Upon completion of the microbiological analyses, the pH values of the vegetable samples were measured with a digital pH meter (RL150, Russell pH Cork, Ireland) with a glass electrode (Metrohm AG, Herisau, Switzerland).

5.2.3. Spectroscopy and multispectral imaging technologies

All the analytical technologies (sensors) used in the present study have been previously described in Chapter 4.

5.2.4. Data analysis

Prior to model development, specific spectral ranges were selected for each one of the analytical technologies. The selected wavenumbers for FTIR were in the range of 1800-870 cm^{-1}, since this range is expected to depict chemical changes related to the metabolic activity of spoilage microorganisms (Al-Jowder et al., 1999; Ellis et al., 2002; Di Egidio et al., 2009). The wavelength range of 700-400 nm was selected for the VIS spectral data, while for NIR all the available spectrum was used in further analyses. As far as MSI is concerned, both the mean reflectance values (18 wavelengths) and their standard deviations (in total 36 features) were used, since standard deviation contains relevant and useful information. All spectral data were normalized under the robust version of Standard Normal Variate (SNV) scheme, aiming at correcting spectra for light scatter and adjusting for baseline shifts among samples (Barnes et al., 1989; Guo et al., 1999). The spectral/imaging data were used as independent variables (X), while the TVC and *Pseudomonas* spp. populations as dependent variables (Y) for the purpose of model development.

For each one of the tested analytical technologies, two data partitioning approaches and two machine learning algorithms were employed and their performance was compared. With reference to data partitioning, the data corresponding to the dynamic temperature conditions were used for external model validation (test set) or, alternatively, a random partitioning of data in training (80%) and test (20%) sets was followed. The machine learning algorithms evaluated herein were: (i) the linear partial least squares regression (PLSR) algorithm, since it is considered suitable for spectroscopy datasets (Mehmood et al., 2012; Wold et al., 2001); and (ii) the support vector regression (SVR) algorithm based on the radial basis function (RBF) kernel, which are among the most popular statistical methods (Cortes & Vapnik 1995; Kiala et al., 2016). In the case of the baby spinach stored under active MAP, only PLSR and random data partitioning were performed since there were no data derived from dynamic conditions.

The performance of the developed PLSR and SVR models for each analytical technology was evaluated based on the statistical parameters of slope (α), offset (b), root mean square error (RMSE) and the coefficient of determination (R-squared, R^2) of the linear regression between the predicted and actual (measured) microbial counts.

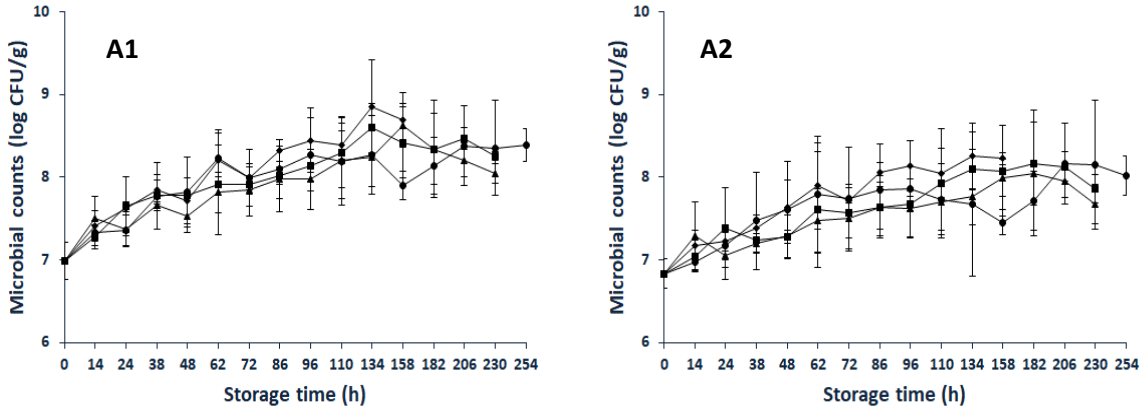
5.3. Results and Discussion

5.3.1. Microbiological spoilage and pH data for vegetables stored under passive MAP

The initial level of TVC (mean ± standard deviation, n=4) for baby spinach was 6.99 ± 0.23 log CFU/g, while the final levels were 8.39 ± 0.20, 8.26 ± 0.05, 8.69 ± 0.32 and 8.04 ± 0.12 log CFU/g at 4 ºC, 8 ºC, 12 ºC and dynamic temperature conditions, respectively (Figure 5.1). The *Pseudomonas* spp. populations had no more than an average of 0.6 log units difference from the TVC throughout storage. On the other hand, the average level of the Enterobacteriaceae family and lactic acid bacteria (LAB) for baby spinach throughout storage and at all storage temperatures was below 6.0 and 4.0 log units, respectively (data not shown). All the above indicate that *Pseudomonas* spp. are the main spoilage microorganisms in baby spinach. With regard to rocket, the initial TVC was 6.63 ± 1.13 log CFU/g, while the TVC recorded at the final stages of storage was 8.03 ± 0.34, 8.70 ± 0.36, 8.81 ± 0.14 and 8.56 ± 0.34 log CFU/g at 4 ºC, 8 ºC, 12 ºC and dynamic temperature conditions, respectively. The average level of LAB did not exceed 5.5 log units throughout storage and at all temperatures, while the members of the Enterobacteriaceae family reached a final average concentration of not more than 7.5 units (data not shown). Consequently, *Pseudomonas* was also identified as the main bacterial genus associated with rocket spoilage, with even smaller differences from the TVC compared to baby spinach. The dominance of *Pseudomonas* in fresh-cut and RTE salads has also been reported by other researchers in the scientific literature (Conte et al., 2008; Tsironi et al., 2017; Giannoglou et al., 2020). Even though the initial TVC and *Pseudomonas* populations were higher in baby spinach, the microbial populations in rocket reached higher final values for all studied temperatures, except for 4 ºC. As expected, the microbial growth was faster at increasing storage temperatures, particularly in the case of rocket as also observed by other investigators (Gu et al., 2018; Giannoglou et al., 2020). Interestingly, the microbial loads of both vegetables were remarkably high even from the beginning of storage. Similar or even higher initial microbial concentrations have been reported by other researchers for both minimally processed produce commodities (Lopez-Velasco et al., 2010; Medina et al., 2012; Rosberg et al., 2020). However, the range between the initial and the final TVC or *Pseudomonas* counts in baby spinach is considerably narrow, namely less than 1.5 log CFU/g. For rocket, the initial and final values lie in an average range of more than 2 log CFU/g, except

for samples stored at 4 ºC. Giannoglou et al. (2020) and Rosberg et al. (2020) have also observed this microbiological range in minimally processed rocket and spinach.

A considerably high variation of TVC and *Pseudomonas* counts was observed among the different experimental replications (different batches) of vegetables but also between the two biological replicates (duplicate samples originating from different packages), in both vegetables but mainly in rocket (Figure 5.1). Indeed, the large variation in microbial populations among the different batches can be attributed to a range of pre-harvest factors, such as genotype, maturity stage, growing and harvesting conditions, but also to post-harvest and pre-processing operations (e.g., washing, drying and packaging) (Martınez-Sanchez et al., 2006; Conte et al., 2008; Medina et al., 2012; Garrido et al., 2015). On the other hand, the differences observed between the duplicate samples tested at each sampling interval within a certain batch can only be regarded as indicative of the extensive biological variability characterizing the microbiota of fresh produce (Rico et al., 2007). It has been opined that produce commodities are also subject to within plant variability (spur age, crop load, plant age and location). Consequently, the great variation in microbial counts observed within same-brand products with identical "use by" dates tested on the same day, may be justifiable for such reasons too (Valentin-Bon et al., 2008).



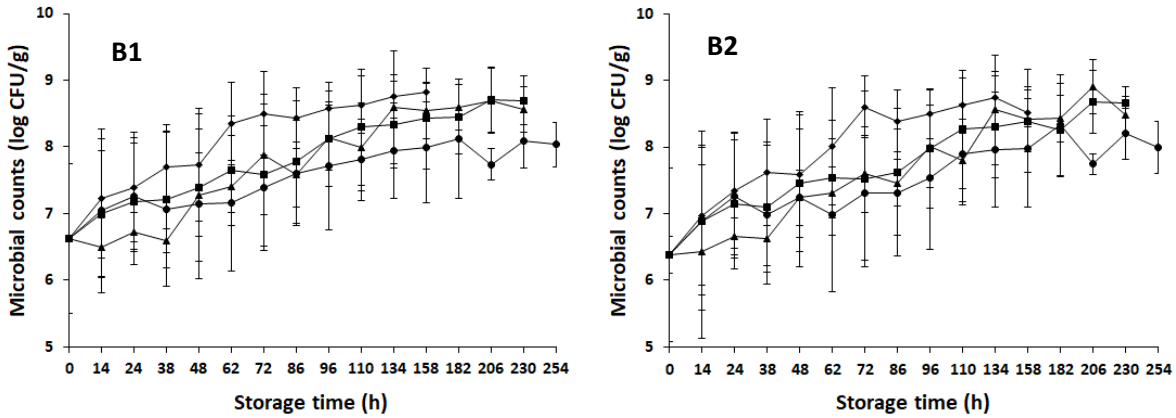


Figure 5.1. Total viable count (TVC) (A1), *Pseudomonas* spp. populations (A2) in ready-to-eat (RTE) baby spinach and TVC (B1), *Pseudomonas* spp. populations (B2) in RTE rocket during storage at 4 ºC (●), 8 ºC (■), 12 ºC (◆), and dynamic temperature conditions (▲) stored under passive modified atmosphere packaging. The populations are expressed as means ± standard deviations (n=4 for baby spinach and n=6 for rocket).

With reference to the pH of the studied vegetables, although the initial values were similar between the two RTE products, the pH of baby spinach finally reached higher values than those recorded for rocket, at all four different temperature conditions (Figure 5.2). A considerable pH increase was observed during baby spinach storage at 12 °C potentially associated with the proteolytic activity of bacterial species belonging to the genus *Pseudomonas*. Nonetheless, no remarkable differences were overall observed in the pH values of rocket among the different temperatures throughout storage.

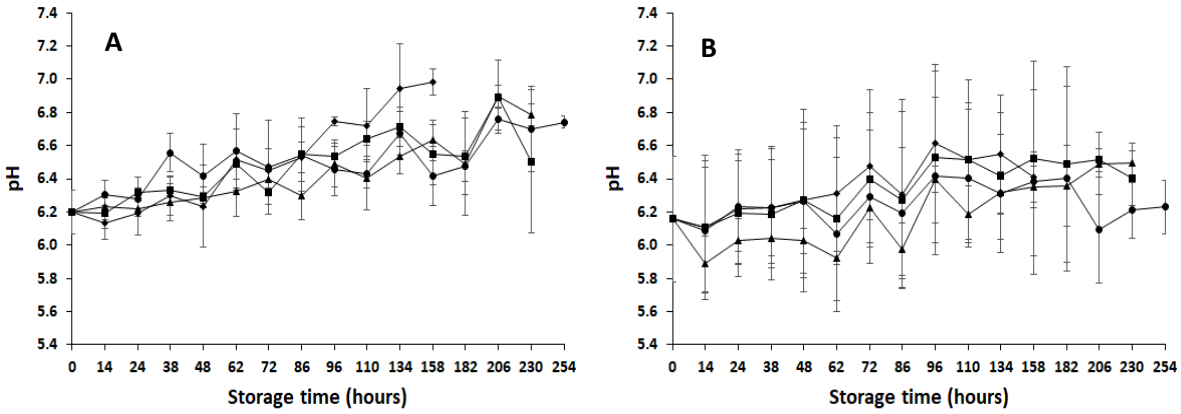


Figure 5.2. The pH values of ready-to-eat baby spinach (A) and rocket (B) during storage at 4 ºC (●), 8 ºC (■), 12 ºC (◆), and dynamic temperature conditions (▲) stored under passive

modified atmosphere packaging. The pH values are expressed as means ± standard deviations (n=4 for baby spinach and n=6 for rocket).

The changes in atmosphere composition inside the packages of leafy vegetables are also presented in Figure 5.3. The initial O_2 concentration in baby spinach was lower (3%) than that of rocket, while the initial level of CO_2 was higher (1.5%). The final gas compositions in baby spinach reached lower levels of O_2 compared to rocket. On the other hand, the final CO_2 percentages were similar for both vegetables when samples were stored at 8 ^oC and dynamic conditions, higher (1.3%) in baby spinach for storage at 4 ^oC, but also higher (2.3%) in rocket at 12 ^oC. The storage temperature influenced considerably the respiration rate. The consumption and production rate of O_2 and CO_2 respectively, was higher at increasing temperatures. Although, rocket showed a higher microbial growth rate compared to baby spinach, the respiration increase of the latter was more intense. This high respiration rate for baby spinach has been also reported in the literature (Allende et al., 2004).

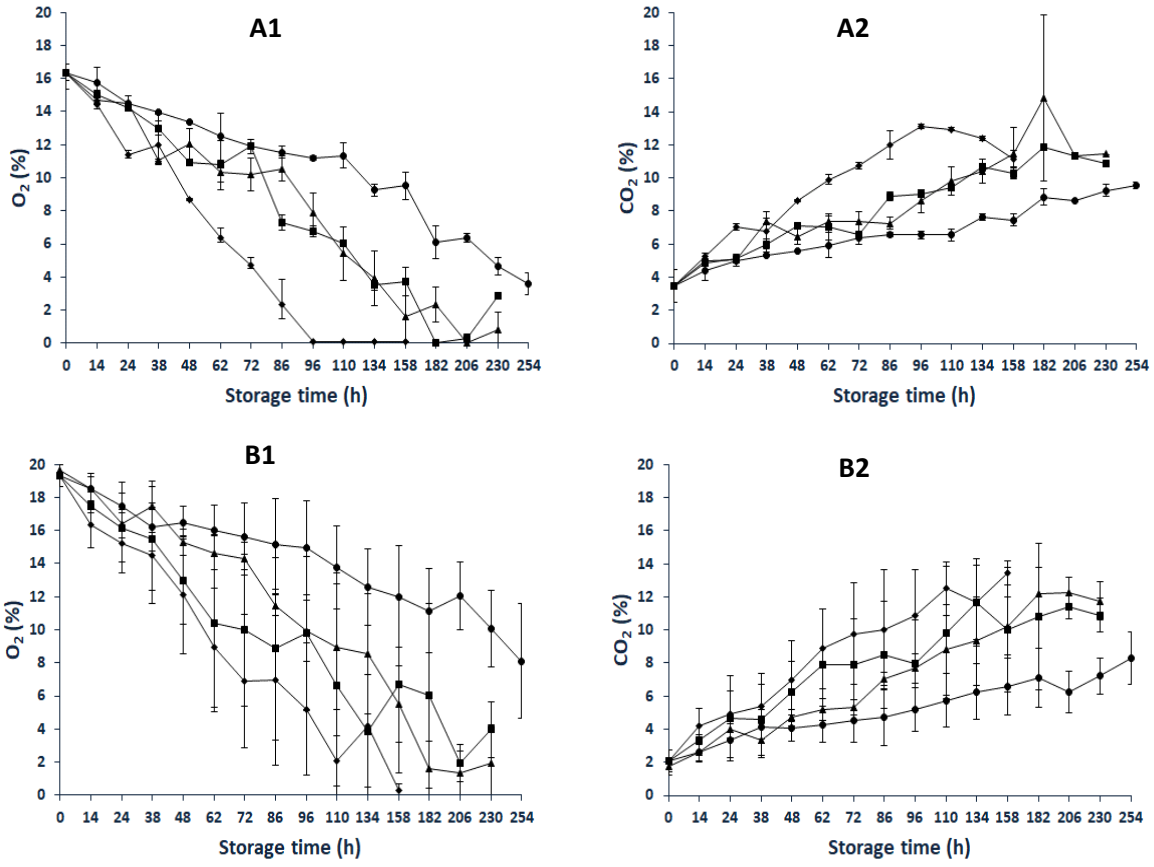


Figure 5.3. O_2 (A1) and CO_2 (A2) compositions in ready-to-eat (RTE) baby spinach, O_2 (B1) and CO_2 (B2) compositions in RTE rocket during storage at 4 ºC (●), 8 ºC (■), 12 ºC (◆), and dynamic temperature conditions (▲) stored under passive modified atmosphere packaging.

For baby spinach, the gas compositions concern only one batch due to experimental inability to measure the corresponding compositions for the second batch and are expressed as mean ± standard deviation (n=2 for duplicate samples). For rocket, the results concern the three batches for the isothermal temperatures, but the two batches for dynamic temperature since there were no samples stored at such conditions for one of the batches The results are expressed as mean ± standard deviation (n=6 for constant temperatures or n=4 for dynamic conditions).

5.3.2. Microbiological spoilage and pH data for vegetables stored under active MAP

The initial TVC and *Pseudomonas* spp. populations (means ± standard deviations, n=2) for baby spinach stored under active MAP was 7.37 ± 0.60 and 7.03 ± 0.60 log CFU/g, respectively (Figure 5.4). The initial microbial counts were not considerably different (i.e. no more than 0.4 log units) between samples stored under active and passive MAP. However, the TVC, but mainly *Pseudomonas* spp. growth, was higher in active MAP at the higher storage temperatures, reaching almost 1-log unit difference compared to passive MAP. Nonetheless, the growth curves recorded at 4 ºC were similar for both packaging conditions. Again, the microbial growth at 12 ºC was higher compared to 4 ºC, while considerable variations in microbial populations among the batches and between the duplicate samples were present. As far as the pH is concerned, the initial value was 6.45 ± 0.05, while the final values corresponding to similar final storage times of passive MAP were 7.07±0.06, 7.02±0.00 and 7.04±0.11 for 4, 8, and 12 ºC, respectively. Therefore, the above values were slightly higher than those of passive MAP. According to the changes in gas composition during storage presented in Figure 5.5, the super atmospheric O_2 inside the packages was finally reduced to values not below 70% for all the tested temperatures, while the CO_2 reached values of 18% at the highest temperature.

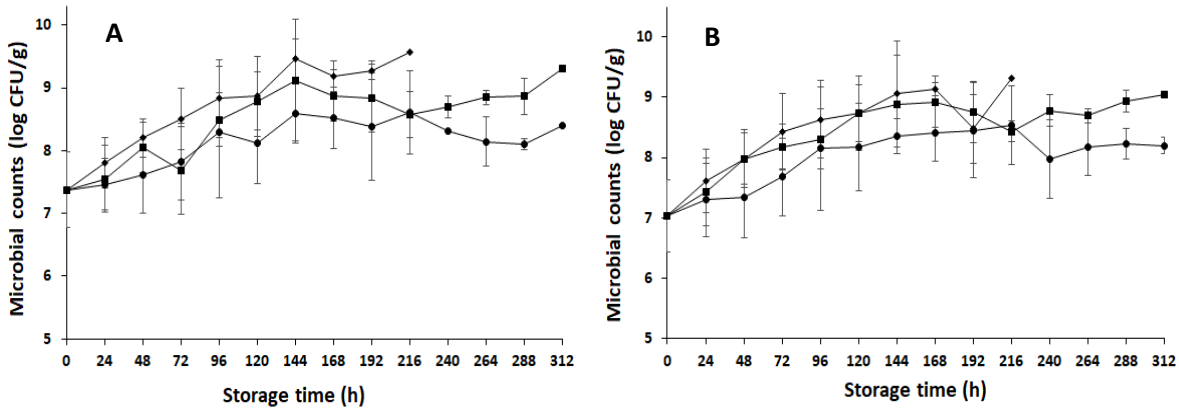


Figure 5.4. Total viable count (A) and *Pseudomonas* spp. populations (B) in ready-to-eat baby spinach stored under active active modified atmosphere packaging at 4 ºC (●), 8 ºC (■) and 12 ºC (◆).The populations are expressed as mean ± standard deviation (n=4).

As it was mentioned above, the temperature influenced the atmosphere composition. Regarding microbial growth, the reports on the impact of super atmospheric O_2 MAP for the different commodities vary considerably in the literature (Oliveira et al., 2015). In the present study microbial growth and mainly the growth of the dominant spoilage group in baby spinach was increased under conditions of super atmospheric O_2. Although the antimicrobial activity of CO_2 at high concentration has been well established, the aerobic populations in the present study increased, although there were higher levels of CO_2 accumulation compared to the passive MAP. Moreover, Allende et al. (2004) reported insignificant reduction in aerobic as well as anaerobic microbial growth in super atmospheric O_2 and perforated MAP compared to the conventional MAP. It should also be mentioned that the sensory quality of baby spinach in the tested active MAP was improved during storage in terms of odour but not in terms of water leakage.

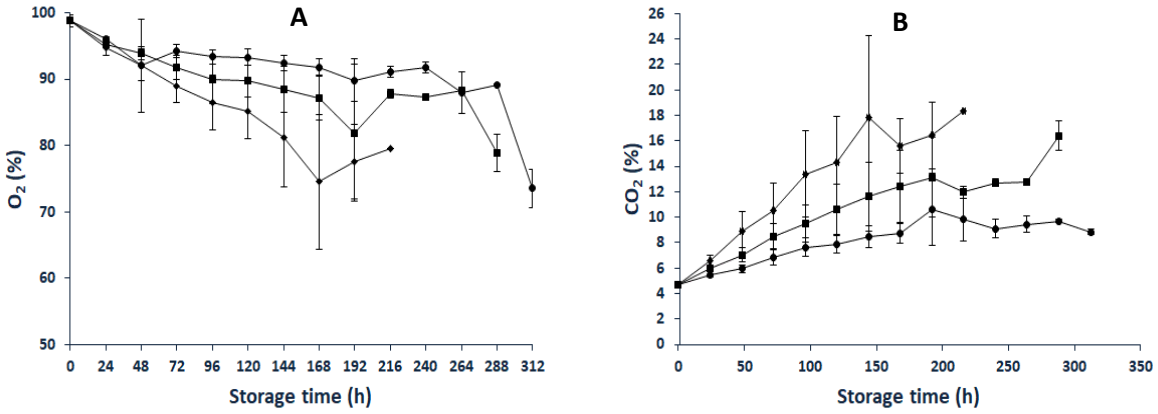


Figure 5.5. O_2 (A) and CO_2 (B) compositions in ready-to-eat baby spinach stored at 4 ºC (●), 8 ºC (■), 12 ºC (◆), as well as under active modified atmosphere packaging. The results are expressed as means ± standard deviations (n=4).

5.3.3. Models for RTE vegetable salads in passive MAP

In the present study, we evaluated not only the prediction of the TVC, but also the *Pseudomonas* spp. populations, since the latter is the dominant spoilage group of microorganisms in RTE vegetable salads. Two approaches of data partitioning into training and external validation sets were applied and evaluated in terms of model performance. Non-isothermal temperatures are simulating better the conditions encountered in the cold chain for refrigerated food products such as RTE vegetables (Ndraha et al., 2018). Hence, the selection of data derived from dynamic temperature conditions as test set, may allow for the development of meaningful predictive models. On the other hand, the random data partitioning approach provides a less biased sample selection, ensuring that all temperature and storage times are equally represented in both training and test datasets, and allowing for robust model training (Manthou et al., 2020). Beyond the above mentioned data partitioning schemes, two machine learning algorithms, namely PLSR and SVR based on the RBF, were employed to unravel which captures better the microbiological spoilage of the tested vegetables. PLSR is a linear approach for modelling the relationship between the dependent and the independent variables, while radial SVR is a non-linear method. The main advantages of SVR is that its computational complexity does not depend on the dimensionality, it has excellent generalization capability and high prediction accuracy (Awad & Khanna, 2015). On the other

hand, PLSR is suitable for spectroscopy datasets that encounter dimensionality issues, strong collinearity and noise (Wold et al., 2001; Mehmood et al., 2012).

5.3.3.1. Baby spinach

In Table 5.1, the prediction performances of PLSR and SVR models for TVC are presented. It was observed that when model testing was performed on samples randomly selected, the performance of the PLSR and SVR models was dependent on the applied analytical technology, being either similar or better compared to the one attained when testing was performed based on the dynamic temperatures data. Specifically, for MSI and NIR sensors the overall performance depicted by the slope, offset and R-squared values was quite better with random partitioning, while the RMSE was low (below 0.5 log CFU/g). For FTIR and VIS, the models exhibited slightly better, but actually in most cases similar overall performances. However, for the combination of FTIR sensor and SVR algorithm, the random partitioning approach resulted in quite higher R-squared value.

Regarding the *Pseudomonas* spp. prediction, relatively similar trends were overall observed for the two data partitioning approaches (Table 5.2). In the case of the VIS and FTIR sensor, the PLSR performance was similar for both approaches. For MSI, higher R-squared value was obtained with the random data partitioning (as compared to the utilization of the dynamic temperatures data set for model external validation), but for NIR the performance was similar for the two approaches. Concerning SVR models, in the case of MSI and NIR the performance metrics (a, b, and R-squared) were notably better in random partitioning, while the RMSE was below 0.5 log CFU/g. For FTIR, the performance was similar for the two approaches. However, when the dynamic data were used as test data set in VIS, the overall performance of the SVR model was better in terms of R-squared value. In general, the random partitioning approach seems to result in better performances for MSI and NIR, but in similar performances when FTIR and VIS are applied.

Table 5.1. Performance metrics of the partial least squares regression (PLSR) and support vector regression (SVR) models for the different analytical technologies (sensors) used in the prediction of total viable count in ready-to-eat baby spinach. Parameters *a*, and *b* are the slope and offset of the linear regression ($y=ax+b$) between predicted and measured

values,respectively, RMSE is the root mean square error of the fit, and R^2 is the coefficient of determination.

	Analytical platform	Algorithm	Data segmentation	*a*	*b*	R^2	RMSE
TVC	FTIR	PLSR	Random	0.54	3.55	0.34	0.45
		SVR	Random	0.48	4.09	0.46	0.35
		PLSR	Dynamic	0.53	3.72	0.28	0.45
		SVR	Dynamic	0.42	4.61	0.32	0.33
	MSI	PLSR	Random	0.64	2.72	0.4	0.51
		SVR	Random	0.48	4.03	0.56	0.33
		PLSR	Dynamic	0.39	4.64	0.27	0.26
		SVR	Dynamic	0.26	5.77	0.24	0.17
	VIS	PLSR	Random	0.74	2.04	0.65	0.4
		SVR	Random	0.69	2.48	0.59	0.39
		PLSR	Dynamic	0.69	2.41	0.57	0.41
		SVR	Dynamic	0.61	3.06	0.61	0.35
	NIR	PLSR	Random	0.63	2.92	0.61	0.41
		SVR	Random	0.55	3.54	0.53	0.39
		PLSR	Dynamic	0.43	4.63	0.41	0.3
		SVR	Dynamic	0.44	4.59	0.39	0.3

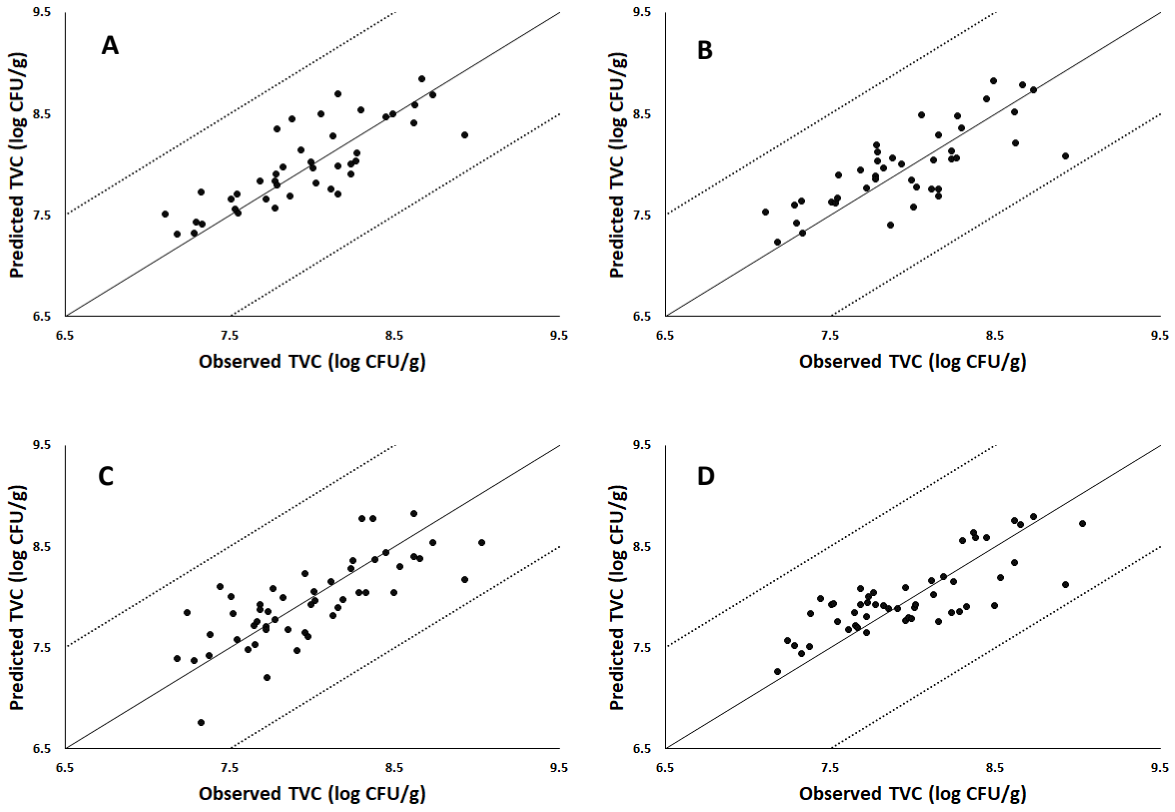
As far as the different regression algorithms employed are concerned, i.e. PLSR and SVR, the resulting models exhibited similar performances in the majority of the cases in terms of TVC prediction. However, the SVR models derived from FTIR and MSI data combined with random data partitioning exhibited quite higher R-squared values. For *Pseudomonas* spp. prediction, the two algorithms resulted also in similar results for the most of the cases. One exception was observed for the VIS sensor, where the SVR model combined with the dynamic partitioning approach showed considerably better slope, offset and R-squared values than the PLSR models. Furthermore, the PLSR model derived from the MSI data and for the dynamic partitioning had considerably higher predictive power than that of the corresponding SVR model.

Table 5.2. Performance metrics of the partial least squares regression (PLSR) and support vector regression (SVR) models for the different analytical technologies (sensors) used in the prediction of *Pseudomonas* spp. populations in ready-to-eat baby spinach. Parameters *a* and *b* are the slope and offset of the linear regression ($y=ax+b$) between predicted and measured values, respectively, RMSE is the root mean square error of the fit, and R^2 is the coefficient of determination.

	Analytical platform	Algorithm	Data segmentation	*a*	*b*	R^2	RMSE
***Pseudomonas* spp.**	FTIR	PLSR	Random	0.56	3.36	0.4	0.37
		SVR	Random	0.5	3.8	0.45	0.32
		PLSR	Dynamic	0.62	2.88	0.32	0.48
		SVR	Dynamic	0.54	3.55	0.41	0.36
	MSI	PLSR	Random	0.56	3.32	0.6	0.39
		SVR	Random	0.51	3.76	0.72	0.32
		PLSR	Dynamic	0.63	2.76	0.41	0.32
		SVR	Dynamic	0.26	5.47	0.2	0.17
	VIS	PLSR	Random	0.47	4.02	0.37	0.37
		SVR	Random	0.63	2.8	0.54	0.41
		PLSR	Dynamic	0.45	4.15	0.36	0.33
		SVR	Dynamic	0.73	2.06	0.69	0.39
	NIR	PLSR	Random	0.72	2.1	0.52	0.48
		SVR	Random	0.69	2.38	0.5	0.47
		PLSR	Dynamic	0.57	3.41	0.4	0.4
		SVR	Dynamic	0.52	3.75	0.3	0.41

Models with relatively good performance for TVC prediction were the models derived from the combinations VIS/PLSR and VIS/SVR under random partitioning, as well as VIS/PLSR with dynamic partitioning. The former exhibited a slightly better performance in terms of offset and R-squared value. The linear regression $y=ax+b$, i.e. first order polynomial regression between the predicted and the measured microbiological values, is presented in

Figure 5.6. The solid lines are the ideal $y=x$ lines (slope = 1 and offset/bias = 0), while the dashed ones exhibit the "acceptable" microbiological limits (±1 log CFU/g). Based on the slope parameter *a* values, ranging from 0.69 to 0.74, a satisfactory correlation was demonstrated between predicted and actually measured TVC values, which was also suggested by the values of the goodness-of-fit measure of R^2 (0.57-0.65). The RMSE values were low from a food microbiology perspective, since deviations in TVC of approximately 0.5 log cycles are rather common (even within the same laboratory using traditional analytical approaches) and justifiable given the extensive variability encountered in biological systems. Parameter values close to the above were obtained for the models developed with the sensor/algorithm combination VIS/SVR under dynamic partition and NIR/PLSR under the random data partitioning scheme (Figure 5.6).



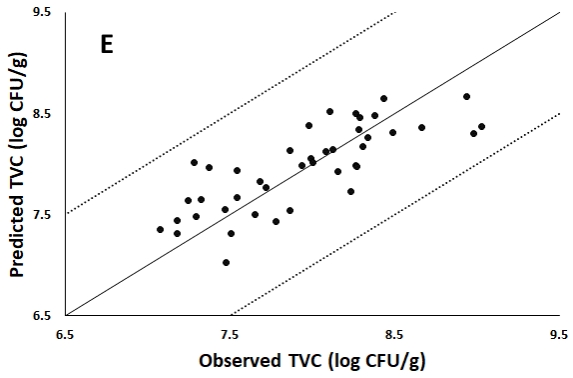


Figure 5.6. Prediction of total viable count in ready-to-eat baby spinach salad: VIS/PLSR with random data partitioning (A), VIS/SVR with random data partitioning (B), VIS/PLSR with dynamic data partitioning (C),VIS/SVR with dynamic data partitioning (D), NIR/PLSR with random data partitioning (E).

Concerning the *Pseudomonas* spp. prediction, the combination VIS/SVR using as test dataset the data from the dynamic temperature conditions exhibited the best performance among the generated models, which was a relatively good performance. The slope value was 0.73, showing good correlation between predicted and actual values, while the R^2 was 0.69. The model derived from the combination of MSI/PLSR using the random approach exhibited performance lower to that of the aforementioned model. Additionally, the models derived from the combinations of NIR/PLSR and NIR/SVR using the random data partitioning scheme showed predictive power less satisfactory, mainly in terms of R-squared value. The slope values of the regression between estimated and measured *Pseudomonas* spp. counts were in the range of 0.69 to 0.72. Additionally, the R-squared values varied from 0.50 to 0.52 and the RMSE values were quite low. The linear first order polynomial regression between the predicted and the observed *Pseudomonas* spp. values of the aforementioned models is presented in Figure 5.7.

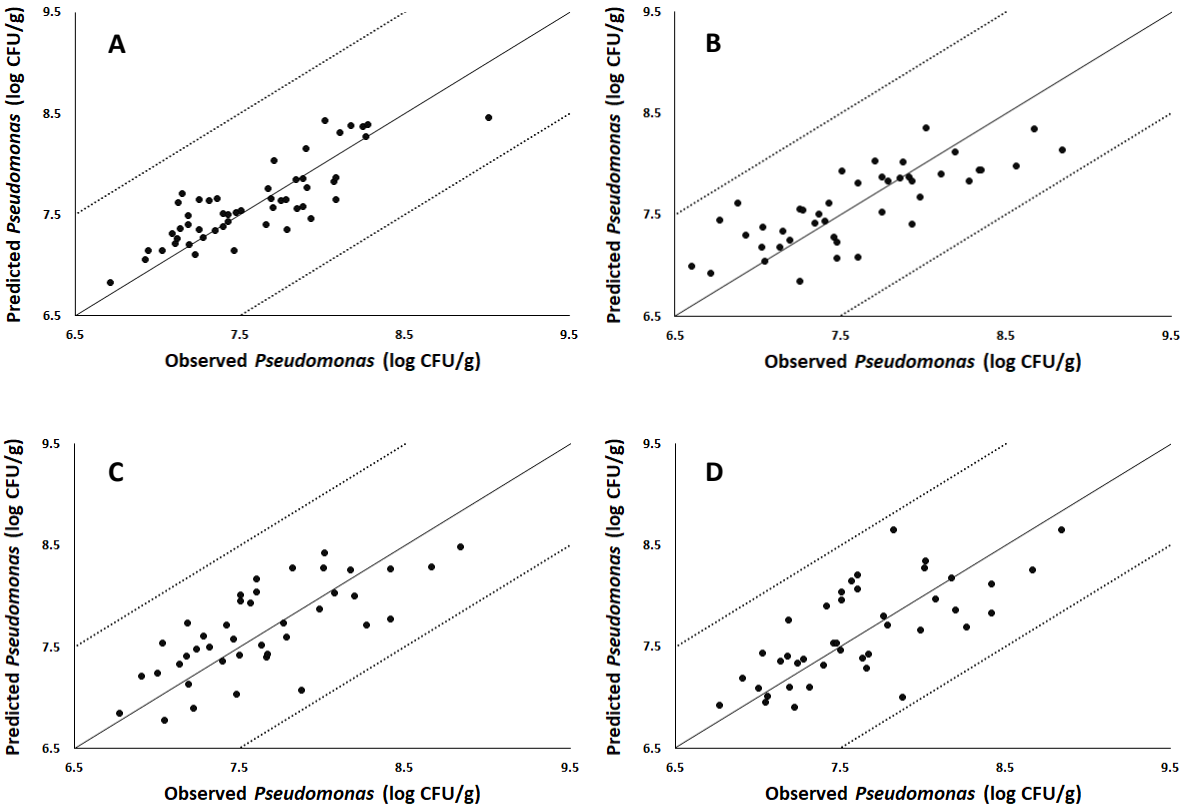


Figure 5.7. Prediction of *Pseudomonas* spp. populations in ready-to-eat baby spinach salad: VIS/SVR with dynamic data partitioning (A), MSI/PLSR with random data partitioning (B), NIR/PLSR with random data partitioning (C), NIR/SVR with random data partitioning (D).

5.3.3.2. Rocket

The prediction performances of the PLSR and SVR models developed under the two data partitioning approaches are presented in Tables 5.3 and 5.4 for TVC and *Pseudomonas* spp., respectively. Concerning the PLSR and SVR models for TVC prediction, the application of random partitioning of the data, resulted in considerably higher predictive power for all the analytical technologies. All the models had better performance in terms of slope, offset, R-squared, while also the RMSE was below 0.8 log CFU/g, which is an acceptable value from a food microbiology perspective. One exception was observed for FTIR and PLSR combination, where the performances for the two different data partitioning approaches were similar. Similar trends were observed for the PLSR and SVR models for *Pseudomonas* spp. prediction, since the performance metrics were considerably better compared to the ones attained using the

dynamic temperature data partitioning, and the RMSE values were also in the acceptable range. The low efficacy of the dynamic temperature data partitioning in rocket compared to baby spinach may be explained by the absence of dynamic conditions for the third batch of rocket. As it was mentioned before, one of the batches of rocket was stored only at isothermal temperatures and therefore, the prediction dataset did not include the variability of all the batches. Additionally, this is supported by the fact that apart from the high variability in microbial counts, we observed great spectral variability among the different batches of rocket (data not shown). Zhang et al. (2018) reported that factors affecting microbial counts in vegetables (e.g., genotype, maturity stage, growing and harvesting environmental conditions, as well as post-harvest and pre-processing operations) have been associated with spectral variability in non-destructive analytical approaches utilizing imaging and NIR spectroscopy technologies. Therefore, the success of random data split in rocket probably lies in the fact that the overall high variability is represented better in this approach for both training and prediction (test) processes.

The comparison of the two algorithms, showed similar trends for both microbial groups (TVC and *Pseudomonas* spp.). Specifically, SVR prevailed considerably or slightly in overall prediction power for FTIR, VIS and NIR sensor, depending on the data partitioning method. The differences were notably larger for the random splitting in most cases, but lower in the dynamic temperature data partitioning, resulting in more or less similar performances. Even though the models validated with the dynamic temperature data for VIS and NIR were not at all appropriate for prediction, the same trend for the two algorithms was observed. However, similar performances of the two algorithms were only observed when random partitioning was used for NIR and FTIR data, in the case of TVC and *Pseudomonas* spp., respectively. On the other hand, PLSR was the best algorithm for the MSI sensor with higher differences for the random but slight differences (similar performance) for the dynamic partitioning approach.

Table 5.3. Performance metrics of the partial least squares regression (PLSR) and support vector regression (SVR) models for the different analytical technologies (sensors) used in the prediction of total viable count in ready-to-eat rocket. Parameters *a* and *b* are the slope and offset of the linear regression ($y=ax+b$) between predicted and measured values, respectively, RMSE is the root mean square error of the fit, and R^2 is the coefficient of determination.

	Analytical platform	Algorithm	Data segmentation	*a*	*b*	R^2	RMSE
TVC	FTIR	PLSR	Random	0.62	2.85	0.41	0.82
		SVR	Random	0.71	2.14	0.70	0.72
		PLSR	Dynamic	0.49	3.91	0.36	0.72
		SVR	Dynamic	0.45	4.10	0.43	0.61
	MSI	PLSR	Random	0.72	2.17	0.61	0.81
		SVR	Random	0.51	3.80	0.59	0.58
		PLSR	Dynamic	0.43	4.13	0.42	0.57
		SVR	Dynamic	0.30	5.07	0.31	0.47
	VIS	PLSR	Random	0.32	5.28	0.17	0.73
		SVR	Random	0.57	3.25	0.56	0.71
		PLSR	Dynamic	0.06	7.16	0.00	1.30
		SVR	Dynamic	0.16	6.24	0.07	0.53
	NIR	PLSR	Random	0.41	4.45	0.31	0.69
		SVR	Random	0.47	4.06	0.43	0.66
		PLSR	Dynamic	0.07	6.94	0.02	0.42
		SVR	Dynamic	0.07	6.90	0.01	0.55

Table 5.4. Performance metrics of the partial least squares regression (PLSR) and support vector regression (SVR) models for the different analytical technologies (sensors) used in the prediction of *Pseudomonas* spp. populations in ready-to-eat rocket. Parameters *a* and *b* are the slope and offset of the linear regression ($y=ax+b$) between predicted and measured values, respectively, RMSE is the root mean square error of the fit, and R^2 is the coefficient of determination.

	Analytical platform	Algorithm	Data segmentation	*a*	*b*	R-squared	RMSE
***Pseudomonas* spp.**	FTIR	PLSR	Random	0.57	3.16	0.48	0.78
		SVR	Random	0.64	2.67	0.61	0.77
		PLSR	Dynamic	0.50	3.77	0.39	0.72
		SVR	Dynamic	0.45	4.08	0.44	0.61
	MSI	PLSR	Random	0.76	1.86	0.75	0.83
		SVR	Random	0.55	3.51	0.62	0.67
		PLSR	Dynamic	0.53	3.21	0.51	0.64
		SVR	Dynamic	0.38	4.38	0.42	0.51
	VIS	PLSR	Random	0.25	5.66	0.33	0.42
		SVR	Random	0.54	3.53	0.50	0.73
		PLSR	Dynamic	-0.09	8.32	0.00	1.33
		SVR	Dynamic	0.14	6.37	0.04	0.59
	NIR	PLSR	Random	0.32	5.07	0.24	0.63
		SVR	Random	0.41	4.55	0.44	0.58
		PLSR	Dynamic	0.14	6.21	0.04	0.62
		SVR	Dynamic	0.11	6.55	0.02	0.57

The best models for TVC prediction were those derived from the combinations FTIR/SVR and MSI/PLSR using the random data partitioning scheme. The former exhibited a slightly better performance in terms of R^2 and RMSE. The slope, offset and R^2 values were 0.71, 2.14, and 0.70, respectively, while the RMSE was 0.72 log CFU/g. The statistical metrics indicated a relatively good performance. The linear first order polynomial regression between

the predicted and the measured TVC values is presented in Figure 5.8. Lower to the above was the performance for the combination VIS/SVR with random data splitting, and even less satisfactory was the predictive power for MSI/SVR using the random approach (Figure 5.8).

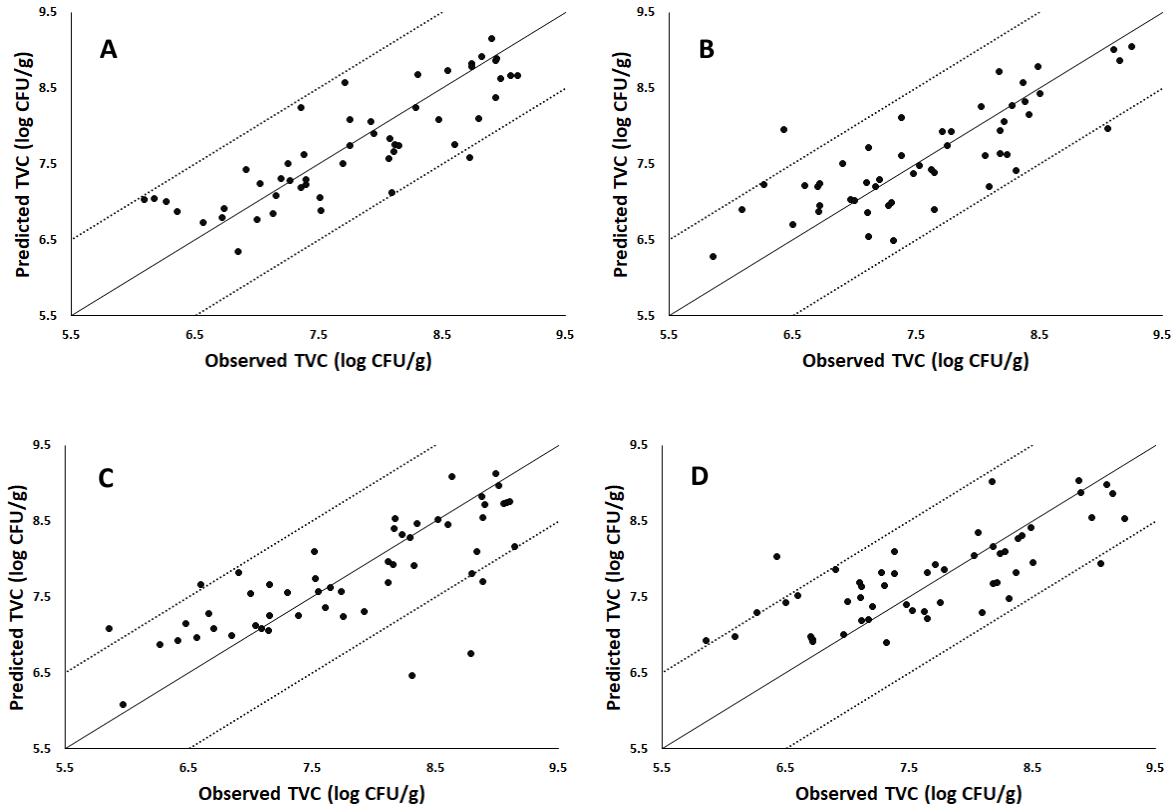


Figure 5.8. Prediction of total viable count in ready-to-eat rocket salad: FTIR/SVR with random data partitioning (A), MSI/PLSR with random data partitioning (B), VIS/SVR with random data partitioning (C), MSI/SVR with random data partitioning (D).

Accordingly, a good model for *Pseudomonas* spp. was also derived from the combination of MSI/PLSR with random data partitioning. In this case, the performance values of slope, offset and R-squared were 0.76, 1.86 and 0.75, respectively and are illustrated in Figure 5.9. The combinations of FTIR/SVR and MSI/SVR with random partitioning exhibited quite lower performances, in terms of R-squared values. Additionally, the combinations of FTIR/PLSR and VIS/SVR with the random approach as well as, MSI/PLSR with the dynamic approach exhibited even lower performances (Figure 5.9).

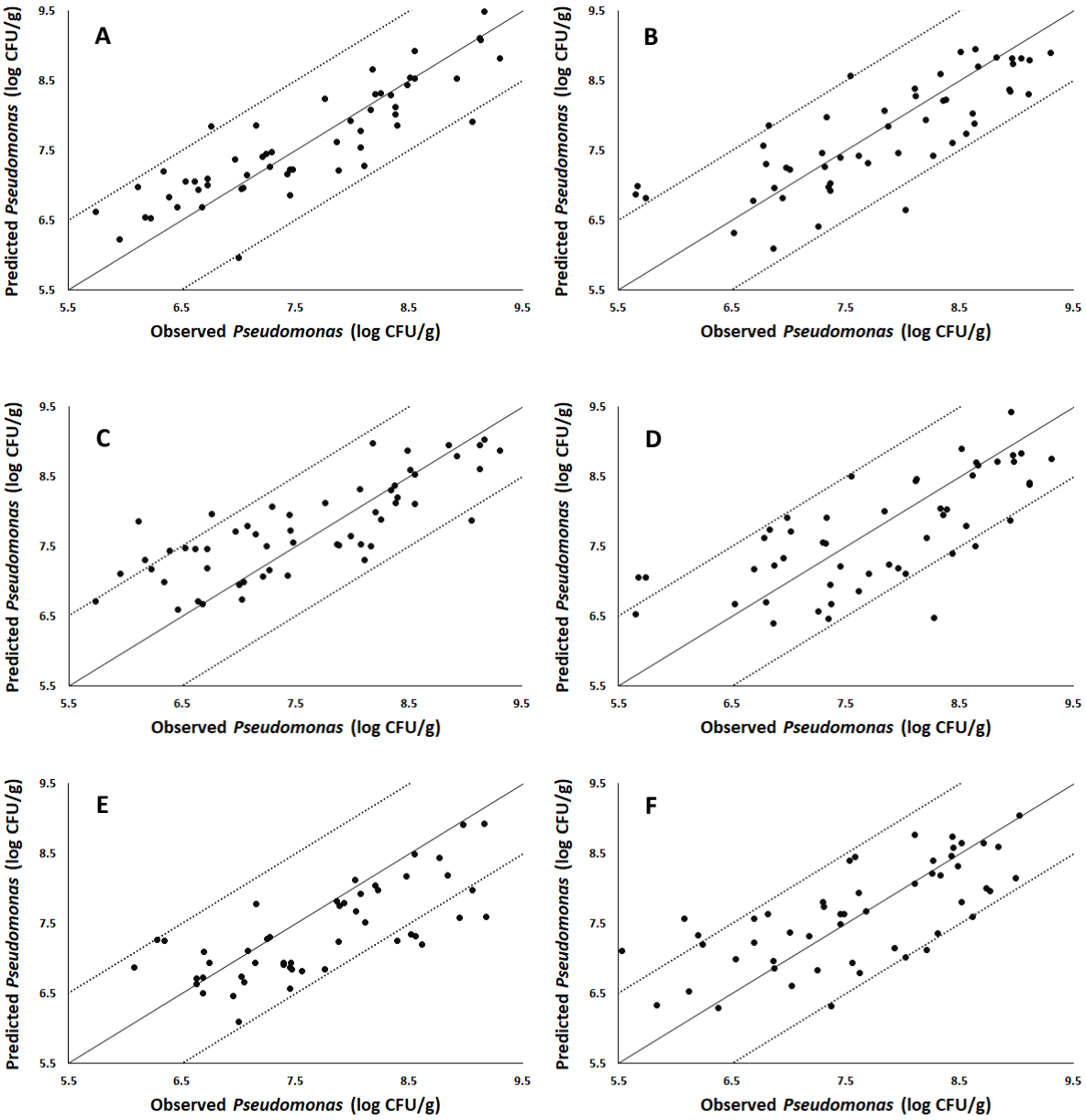


Figure 5.9. Prediction of *Pseudomonas* spp. populations in ready-to-eat rocket salad: MSI/PLSR with random data partitioning (A), FTIR/SVR with random data partitioning (B), MSI/SVR with random data partitioning (C), FTIR/PLSR with random data partitioning (D), MSI/PLSR with dynamic data partitioning (E), VIS/SVR with random data partitioning (F).

Interestingly, FTIR and MSI sensors seem to be more appropriate for rocket, while mainly VIS but also NIR were evaluated as such for baby spinach. According to literature, NIR, as well as VIS/NIR and MIR (FTIR) are the regions that contain the greatest number of applications pertaining to fruit and vegetable. Mainly, NIR and FTIR spectroscopy have been successfully applied in quality evaluation of vegetables (Bureau et al., 2019; Krivoshiev et al.,

2000; Nicolai et al., 2007). More recently, multispectral imaging methods have been used to determine the quality of various fruits and vegetables (Løkke et al., 2013; Lunadei et al., 2012). However, the application of the aforementioned analytical technologies in evaluation of vegetable microbiological spoilage, has been studied only for shredded cabbage, using NIR spectroscopy (Suthiluk et al., 2008).

5.3.4. Models for RTE baby spinach in active MAP

The prediction performances of PLSR models for TVC estimation of RTE baby spinach stored under active MAP are presented in Table 5.5. Additionally, the linear first order polynomial regression between the predicted and the measured TVC values is illustrated in Figure 5.10. The populations of the dominant spoilage group, namely *Pseudomonas* spp., were not evaluated herein, as the results in subsection 5.3.2. showed the same trends for both microbial groups. The presented models were developed using only random data partitioning due to the lack of samples stored under dynamic temperature conditions. However, the random partitioning scheme exhibited similar or better performance compared to dynamic partitioning in baby spinach stored under passive MAP, indicating that the former is an ideal option in any case. Moreover, the PLSR algorithm was the only algorithm applied in the present dataset since SVR showed similar results to PLSR for all the developed models in the passive MAP case. With regard to the performance metrics, both models derived from FTIR and MSI data exhibited good performances, similar to those of the best models presented for baby spinach stored under passive MAP in terms of $R^2$and RMSE. However, the values of slope and offset showed a slightly better predictive power. Interestingly, FTIR and MSI were appropriate sensors for TVC prediction of the baby spinach samples stored under active MAP but not for the samples stored under passive MAP. Therefore, the sensors' applicability may be not only product-specific, but also storage-specific.

Table 5.5. Performance metrics of the partial least squares regression (PLSR) models for the different analytical technologies (sensors) used in the prediction of total viable count in ready-to-eat baby spinach stored under active modified atmosphere packaging. Parameters *a* and *b* are the slope and offset of the linear regression ($y=ax+b$) between predicted and measured

values, respectively, r is the correlation coefficient, RMSE is the root mean square error of the fit, and R^2 is the coefficient of determination.

Sensors	*a*	*b*	r	R^2	RMSE
FTIR	0.83	1.22	0.84	0.57	0.48
MSI	0.85	1.13	0.85	0.67	0.42

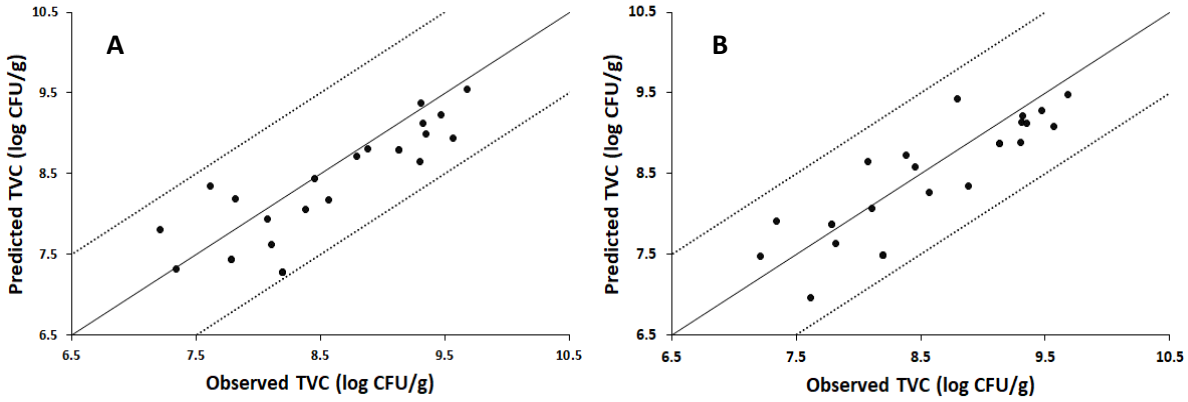


Figure 5.10. Comparison between the observed and the predicted total viable count by the partial least squares regression model based on FTIR (A) and MSI (B) data of ready-to-eat baby spinach stored under active modified atmosphere packaging for the external validation dataset (solid line: the ideal y=x line; dashed lines: the ± 1 log unit area).

5.4. Conclusions

In the present study, for the first time, to our knowledge, different data partitioning schemes and machine learning algorithms were comparatively evaluated for assessing the microbiological quality of vegetable products using various spectroscopy-based technologies. The results obtained in this work demonstrate that the random data partitioning scheme resulted in considerably better performance in almost all cases of sensor/algorithm combination of rocket. On the other hand, it was observed that when random partitioning was selected for baby spinach, the performance of the models was often dependent on the applied analytical technology, being either similar (for FTIR and VIS) or better (for NIR and MSI) compared to the one attained when testing was performed based on the dynamic temperatures data. With

regard to the different algorithms used, the predictive power was similar for the two algorithms in the majority of baby spinach cases. In rocket, SVR algorithm provided considerably or slightly better models for the FTIR, VIS and NIR sensors, depending on the data partitioning method. The differences were notably larger for the random approach in most cases, but lower in the dynamic temperature data partitioning, resulting in similar performances. Contrarily, PLSR was the best algorithm for the MSI sensor with considerably higher differences for the random but slight differences (similar performance) for the dynamic partitioning approach.

The present study also revealed that mainly VIS spectroscopy could be more appropriate for TVC and *Pseudomonas* spp. population prediction in RTE baby spinach, while FTIR spectroscopy and MSI for RTE rocket. Although both salad products belong to the leafy vegetables category and seem quite similar to each other, it is evident that they should be treated distinctively in terms of sensor and computational analysis application.

With regard to the baby spinach stored under active MAP, FTIR and MSI sensors were interestingly appropriate for microbiological spoilage evaluation. Therefore, the application of the various sensors may also differ for the different packaging conditions of the same product.

Future research should, by all means, contribute to improved characterization of the extensive physical and biological variability characterizing fresh produce commodities, through the administration of large-scale experiments involving RTE fresh salads such as the ones studied herein. Moreover, additional computational analysis information including, among others, selection of the most informative (in each case) wavelengths and data fusion approaches, should also be at the disposal of researchers, allowing for the development of reliable and robust prediction models and contributing to the expansion of the current knowledge.

Acknowledgments

This work has been supported by the project “PhasmaFOOD”, funded from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 732541.

Chapter 6

Assessment of microbiological spoilage of fresh-cut oyster mushrooms through Fourier transform infrared spectroscopy and multispectral imaging

Chapter 6

Abstract

Pleurotus ostreatus (oyster mushroom) is one of the most commercially important edible mushrooms worldwide. However, the limited shelf-life of this produce commodity requires rapid and efficient methods for assuring its quality. The aim of this study was the assessment of the microbiological spoilage of oyster mushrooms under different temperature conditions, using Fourier transform infrared (FTIR) spectroscopy and multispectral imaging (MSI). Oyster mushrooms were stored at isothermal (4, 10 and 16 ^{o}C), and dynamic temperature (12 h at 7 °C and 12 h at 13 °C) conditions. At regular time intervals during storage, duplicate samples were analyzed for the determination of the total mesophilic microbial populations, while, in parallel, spectral measurements and multispectral images were acquired for both the cap (pileus) and the gills (lamellae) side of the mushrooms. Three independent experiments were conducted and partial least squares regression (PLSR) was applied to establish the correlation between spectral and microbiological data. Different data partitioning approaches into training and validation sets, as well as selection of the most informative independent variables (spectra) were tested in order to achieve better model performances. Moreover, given the high variability observed between duplicate samples, for both the microbial counts and the spectral measurements, the average values were used for model development. The models derived from FTIR and MSI data exhibited poor predictive power for both mushroom sides. The results indicated that, under the conditions of this study and for the applied data analysis, the application of FTIR and MSI is not promising for the evaluation of the microbiological spoilage of oyster mushrooms.

6.1. Introduction

In recent years, there has been an ongoing increase in the worldwide production and consumption of fresh edible mushrooms (Venturini et al., 2011). Mushrooms are appreciated not only for their unique flavour and texture, but also for their nutritional value and functional properties (Jafri et al., 2013). They constitute an important source of high-quality protein, dietary fibers, vitamins and minerals, while they are low in sodium, fat and cholesterol (Jayawardena & Silva 2013; Ventura-Aguilar et al., 2017). These products are also rich in various nutraceutical compounds, including β-glucans, specific unsaturated fatty acids, triterpene and antioxidant compounds (Meenu & Hu, 2019). Among all edible species, the oyster mushroom (*Pleurotus ostreatus*), is the second most cultivated mushroom worldwide, following *Agaricus bisporus* (Ding et al., 2011).

Oyster mushrooms, similar to other species of mushroom-producing fungi (mainly belonging to the Phylum Basidiomycota), present increased respiration, metabolic rate and several biochemical reactions after harvest, resulting in rapid quality deterioration. They are also an ideal medium for microbial growth, since they have high moisture content, a water activity of 0.98 or higher and neutral pH. The high initial microbial load accelerates mushrooms' spoilage and limits their post-harvest shelf-life to a few days (Venturini et al., 2011; Sun & Li, 2017). Consequently, the development of rapid and effective spoilage detection methods are important for their quality assurance, as well as the reduction of food waste and economical losses for the food industry (Wang et al., 2018).

Although conventional microbiology (e.g. colony counting methods) is commonly used for monitoring of food spoilage, the scientists have to deal with important restrictions. Specifically, the culture-based techniques are time-consuming providing retrospective results, costly and destructive to test products (Nychas et al., 2016). On the other hand, spectroscopy-based analytical technologies have been forwarded for rapid and non-destructive food spoilage evaluation. Compared to conventional methods, they require minimal sample preparation, and they are cost-effective and appropriate for in-, on- and at-line detection of quality/safety issues (Efenberger-Szmechtyk et al., 2018). However, the interpretation of the enormous amount of data generated by these analytical technologies is a very challenging task. Depending on the type of sensor and data complexity, the data analysis involves advanced computational methods, including machine learning (Ropodi et al., 2016). Therefore, various machine

learning approaches coupled with spectroscopy-based technologies have been proposed for the qualitative and/or quantitative evaluation of the microbiological quality and safety of various food products including meat and fish products, dairy products, fruits and vegetables (Al-Holy et al., 2015; Barbin et al., 2015; Sravan-Kumar et al., 2015; Tsakanikas et al., 2018; Fengou et al., 2019a, 2019b; Lianou et al., 2019; Manthou et al., 2020; Spyrelli et al. 2020).

Among the various spectroscopy-based technologies that have been implemented in a broad range of food products, Fourier transform infrared (FTIR) spectroscopy is the most widely applied technology in fresh mushrooms. Near-infrared (NIR) spectroscopy and hyperspectral imaging have also been reported in the literature (Taghizadeh et al., 2009; Gaston et al., 2010; Taghizadeh et al., 2010; Meenu & Hu, 2019). However, the majority of these technologies has been mainly used to explore *Ganoderma lucidum* and only a few studies have reported their application in *Pleurotus* spp. (Meenu & Hu, 2019). Moreover, to date, these technologies have been only proposed to identify various filamentous fungi, discriminate the species according to their geographical origin, assess potential adulterations and evaluate the post-harvest quality and chemical composition (Bekiaris et al., 2020).

Therefore, given the above and the practical absence of relevant data, the objective of the present study was to assess the microbiological spoilage of oyster mushrooms using FTIR spectroscopy and multispectral imaging (MSI) technologies.

6.2. Material and Methods

6.2.1. Sample preparation and storage conditions

Oyster mushrooms were obtained from a local producer (Athens, Greece) and transferred to the laboratory under controlled refrigeration temperature (*ca.* 4 °C) conditions within 24 h from harvest. Four or five mushroom portions (consisting of both the cap/gills and the stipe) were placed in styrofoam trays (22.5×13.5×3 cm), and were covered with household-use cling wrap (cellophane membrane) which was perforated at several points to allow for aerobic storage conditions and unobstructed sample respiration. The trays were stored at different isothermal (4, 10 and 16 °C) and dynamic temperature (12 h at 7 °C and 12 h at 13 °C) conditions in high-precision (±0.5 °C) incubation chambers (MIR-153, Sanyo Electric Co., Osaka, Japan) for a maximum duration of 10 days. The applied storage temperatures were within the temperature range recorded throughout the food supply chain (Giannakourou et al.,

2001), including retail premises (super markets, groceries and street markets), and/or reported in the ComBase database (https://www.combase.cc/index.php/en/). Throughout storage of mushroom samples, the incubation temperatures were recorded at 15-min intervals using electronic temperature-monitoring devices (Cox Tracer, Cox Technologies Inc., Belmont, NC, USA).

On the day of arrival to the laboratory (time-zero) and at regular time intervals during storage, depending on the applied storage temperature, duplicate mushroom samples (i.e. distinct trays) were subjected to: i) microbiological analyses and pH measurements; (ii) FTIR spectroscopic measurements; and (iii) acquisition of multispectral images. The spectroscopy and imaging data were collected for both the cap (pileus) side and the gills (lamellae) side of mushrooms. Three independent experiments, corresponding to different time instances and mushroom batches, were conducted in the context of the present study, and a total of 237 mushroom samples were analysed.

6.2.2. Microbiological analysis and pH measurements

A 25-g portion of mushroom was aseptically weighed in a 400-ml sterile stomacher bag (Seward Medical, London, UK) containing 225 ml of sterile quarter strength Ringer's solution (LAB M Limited, Lancashire, UK), and homogenized in a Stomacher apparatus (Lab Blender 400, Seward Medical) for 60 sec at room temperature. Appropriate serial decimal dilutions in Ringer's solution were surface plated on tryptic glucose yeast agar (Biolife, Milan, Italy) for the determination of total mesophilic microbial populations (total viable count, TVC), and *Pseudomonas* agar base supplemented with CFC (Cephalothin, Fucidin, Cetrimide) selective supplement (LAB M Limited) for the enumeration of *Pseudomonas* spp populations. Total mesophiles and presumptive *Pseudomonas* spp. were enumerated after incubation of the corresponding plates at 25 °C for 72 and 48 h, respectively. The obtained microbiological data were converted to log (colony forming units) per gram of mushrooms (log CFU/g).

Upon completion of the microbiological analyses, the pH values of the mushroom samples were measured using a digital pH meter (RL150, Rusell pH, Cork, Ireland) with a glass electrode (Metrohm AG, Herisau, Switzerland).

6.2.3. FTIR spectroscopy and MSI

The analytical technologies applied in the present study have been previously described in Chapter 4.

6.2.4. Data analysis

Prior to model development, specific spectral ranges were selected for each one of the two analytical technologies. The selected wavenumbers for the FTIR sensor were in the range of 1800-870 cm^{-1}, since this range is expected to depict chemical changes related to the metabolic activity of spoilage microorganisms (Al-Jowder et al., 1999; Ellis et al., 2002; Di Egidio et al., 2009). As far as the MSI is concerned, both the mean reflectance values (18 wavelengths) and their standard deviations (in total 36 features) were used, since standard deviation contains relevant and useful information. Moreover, several pre-processing techniques were tested on each dataset with the aim to minimize any irrelevant information such as noise, particle size deviations, scattering and drifting effects (Wang et al., 2015; Dixit et al., 2017; Suhandy & Yulia, 2017; Li et al., 2018). All the spectral data were mean-centered and scaled (1/SDEV). In the case of FTIR, the data were also pre-processed using the Savitzky–Golay smoothing numerical algorithm with a second-order polynomial and a 9-point window. On the other hand, for MSI, standard normal variate (SNV) was the pre-processing treatments applied for the data corresponding to gills, while no further treatment was used for the data corresponding to caps. The spectral/imaging data were used as independent variables (X), while the TVC as dependent variables (Y) for the purpose of model development.

Partial least squares regression (PLSR) was performed for the correlation between spectral data and TVC. This machine learning algorithm is considered suitable for spectroscopic datasets, where the dimensionality problem exists and also when the data show strong collinearity and noise (Wold et al., 2001; Mehmood et al., 2012; Gromski et al., 2015). The data derived from isothermal storage temperatures were used for the calibration process (training set) and those derived from dynamic temperature conditions for external validation (test set). However, other data partitioning methods were also tested. Specifically, the selection of a whole batch as test set as well as the random partitioning scheme were applied. During the calibration process, leave-one-out cross validation in parallel with Marten's uncertainty test were employed in order to eliminate the risk for over-fitting and test the predictive significance

of the model, but also in order to select the significant X-variables (Westad & Martens, 2000; Wold et al., 2001). Additionally, a feature selection of significant wavenumbers on the basis of random forests (RFs) regression ensemble (Breiman, 2001) was also tested for the FTIR data. Concerning feature selection, spectra were initially normalized under the robust version of SNV, namely RNV (Barnes et al., 1989; Guo et al., 1999), while for model development the random data partitioning was used.

The prediction performance of the developed PLSR models for each sensor was evaluated based on the following statistical parameters: slope (*a*), offset (*b*), correlation coefficient (r), the root mean square error (RMSE), and the coefficient of determination (R^2) of the linear regression between the predicted and measured microbial counts.

6.3. Results and Discussion

6.3.1. Microbiological spoilage and pH data

The TVC values of oyster mushrooms during storage at different temperatures are presented in Table 6.1. The initial level of TVC (mean ± standard deviation, n=6) was 5.24 ± 0.73 log CFU/g, while the final populations were 7.02 ± 0.69, 6.74 ± 0.53, 8.36 ± 0.50 and 7.54 ±0.61 log CFU/g during storage at 4, 10, 16 °C and under dynamic temperature conditions, respectively. Accordingly, the initial and final populations of *Pseudomonas* spp. (mean ± standard deviation, n=6) were similar to TVC, indicating that species of this genus constitute the dominant spoilage microorganisms of oyster mushrooms (data not shown). Indeed, *Pseudomonas* spp. have been strongly associated with mushroom pathology (Largeteau & Savoie, 2010; Sajben et al., 2011) and spoilage (Santana, et al., 2008; Wang et al., 2017). The microbial populations observed in the present study are consistent to those reported in previous studies (Ding et al., 2011; Venturini et al., 2011; Wang et al., 2017). Moreover, as expected, the temperature of storage influenced microbial growth, with the later being faster at higher temperatures.

Interestingly, a considerably high variation of TVC and *Pseudomonas* counts was observed among the different experimental replications (different batches) of mushrooms, but also between the biological replicates (duplicate samples originating from different packages) (data not shown). Various pre-harvest and post-harvest factors may contribute to the observed high inter-batch variability (Martınez-Sanchez et al., 2006; Conte et al., 2008; Medina et al.,

2012; Garrido et al., 2015). With reference to the intra-batch variability, this observation may be attributed to variability within *P. ostreatus* fruiting bodies arising from spatial characteristics during growing period, the abiotic environment's direct effect as well as biotic interactions.

Table 6.1. Total mesophilic microbial populations (total viable count, TVC) on oyster mushrooms during storage at different temperature conditions.

TVC (log CFU/g)[a]						
Storage temperature	**4 °C**			**10 °C**	**16 °C**	**Dynamic**
Storage time (h)	0	5.24 ± 0.73	**Storage time (h)**	0	5.24 ± 0.73	5.24 ± 0.73	5.24 ± 0.73
	38	4.85 ± 1.13		14	5.16 ± 1.01	5.59 ± 0.51	4.85 ± 0.57
	62	5.48 ± 0.66		24	5.69 ± 0.92	5.00 ± 0.62	4.98 ± 1.24
	86	5.27 ± 0.52		38	6.06 ± 0.97	6.17 ± 1.26	5.57 ± 1.32
	110	5.35 ± 0.70		48	6.56 ± 0.44	6.10 ± 0.90	5.89 ± 0.90
	134	6.31 ± 0.36		62	5.82 ± 0.73	6.91 ± 0.83	5.66 ± 0.89
	158	6.52 ± 0.87		72	5.38 ± 0.93	7.22 ± 0.88	6.25 ± 1.20
	182	6.84 ± 0.56		86	6.31 ± 0.73	6.93 ± 1.22	6.12 ± 1.04
	206	6.49 ± 0.35		96	6.68 ± 0.57	7.73 ± 0.61	6.59 ± 0.80
	230	6.87 ± 0.63		110	6.96 ± 0.70	8.16 ± 0.48	6.45 ± 1.32
	254	7.02 ± 0.69		120	6.74 ± 0.53	8.36 ± 0.50	7.54 ± 0.61

[a] Values are means ± standard deviations (n = 6)

With reference to the pH of oyster mushrooms, no remarkable changes were observed during storage at 4 and 10 °C (Table 6.2). However, a considerable pH increase was observed during mushroom storage at 16 °C, potentially associated with the proteolytic activity of

bacterial species belonging to the genus *Pseudomonas* (Franzetti & Scarpellini, 2007; Caldera et al., 2016).

Table 6.2. The pH values (means ± standard deviations, n=6) of fresh-cut mushrooms during storage at different temperatures.

Storage temperature	Initial pH	Final pH
4 ºC	5.92±0.11	5.93 ± 0.04
10 ºC		5.95 ± 0.08
16 ºC		6.52 ± 0.33
Dynamic		5.91 ± 0.08

*The final storage time was 254 h for 4 °C and 120 h for the rest storage temperature conditions.

6.3.2. Models

Different approaches of data partitioning into training and external validation sets were applied and evaluated in terms of model performance. The random data partitioning approach provides a less biased sample selection and ensures that all temperature and storage times are equally represented in both training and testing datasets (Manthou et al., 2020). On the other hand, the selection of dynamic data as test set, may allow for the development of more meaningful predictive models, since the non-isothermal temperatures are more realistically imprinting the fluctuations in the food supply cold chain (Ndraha et al., 2018). Moreover, the selection of data derived from a whole batch as test set was also performed in order to examine more options for achieving the best model performance. Apart from the selection of the most influential X-variables based on Marten's uncertainty, a feature selection scheme based on RFs regression ensemble was also applied in the case of FTIR. The reduction of data dimensionality is an important process for avoiding overfitting and removing features that are not correlated to the phenomenon studied (Tsakanikas et al., 2018).

According to Principal Component Analysis (PCA) plots, a great spectral variability within batches was observed (data not shown). The high spectral variability was also obvious between the duplicate samples originating from different packages and subjected to identical storage conditions. As it was mentioned before, it is well established that agricultural products show high biological variability, resulting in correspondingly high spectral variability

(Esquerre et al., 2012; Zhang et al., 2018). In the context of this observation, averaging of the spectral and microbiological data of duplicate samples was applied. It could be argued that the inclusion of all the variability present is important for developing models with increased generalization properties. However, averaging was practiced herein in order to simplify the process, reduce the variability and reveal potential correlations between spectra and microbial counts.

6.3.2.1. FTIR sensor

The performance parameters of all the developed PLSR models derived from FTIR data are presented in Table 6.3. Regardless of the data partition scheme and the reduction of within-batch variability, all the models derived from FTIR data, for both the caps and gills, presented considerably poor prediction performance. Therefore, the models were not capable to estimate the microbiological spoilage of oyster mushrooms. The reduction of variability by mean calculation of spectra and microbial values of the duplicate samples, seemed to provide slightly better predictions for the majority of cases. The feature selection approach based on RFs was also ineffective in considerably improving the model performance (data not shown).

Previous studies utilizing FTIR spectroscopy for the evaluation of mushroom quality, have an important technical difference with the present study: the mushroom samples were freeze-dried and ground into fine particles (Gorman et al., 2010; Bekiaris et al., 2020). Consequently, the excessive water levels (that potentially mask the relevant information to microbiological spoilage) were reduced and a significantly higher sample homogeneity was achieved compared to the present study. Moreover, the gills, whose physical structure may be challenging for spectra acquisition, were also included in dried samples providing additional information. However, this type of preparation is time consuming, destructive and limits the industrial applicability of FTIR spectroscopy.

Table 6.3. Performance parameters of the partial least squares regression models, based on FTIR data, for the total viable count prediction of fresh mushrooms. (*a*: slope, *b*: offset, r: correlation coefficient, RMSE: root mean square error (log CFU/g), R^2: coefficient of determination).

	Side	Duplicate samples	Test set	*a*	*b*	r	R^2	RMSE
FTIR	Caps	No averaging	Dynamic	0.01	5.95	0.04	-0.08	1.26
			Random	0.2	4.99	0.3	-0.07	1.19
			Batch 1	0.25	4.57	0.4	0.07	1.08
			Batch 2	0.33	4.36	0.36	-0.18	1.03
			Batch 3	0.09	6.09	0.13	-0.62	1.67
		Averaging	Dynamic	0.08	5.49	0.29	0.08	1.1
			Batch 1	0.43	3.26	0.58	0.17	0.93
			Batch 2	0.53	3.01	0.41	-0.62	1.05
			Batch 3	0.14	5.68	0.45	-0.13	1.27
	Gills	No averaging	Dynamic	0.08	5.6	0.21	0.01	1.21
			Random	0.27	4.47	0.47	0.17	1.05
			Batch 1	0.2	5.42	0.33	-0.15	1.27
			Batch 2	0.19	4.93	0.22	-0.46	1.15
			Batch 3	0.23	5.31	0.52	-0.12	1.39
		Averaging	Dynamic	0.17	5.11	0.34	0.07	1.07
			Batch 1	0.43	3.53	0.44	-0.11	1.07
			Batch 2	0.23	4.61	0.25	-0.55	1.03
			Batch 3	0.25	5.08	0.58	-0.04	1.22

6.3.2.2. MSI

The performance parameters of all the developed PLSR models for MSI data are presented in Table 6.4. In the case of caps, the overall performance of all the developed models is not by no means satisfying for mushroom spoilage evaluation. Although a similar trend was observed for the gills side, the model developed after variability reduction and using Batch 1

as test set showed a satisfactory performance. The slope parameter *a* indicated a relatively satisfatory correlation between predicted and actually measured TVC values (0.63), which is also suggested by the goodness-of-fit measure of R^2. It should also be mentioned that the RMSE value was below 1 log CFU/g, which is acceptable for applications in food microbiology. Among all these models with poor prediction efficacy, the relatively satisfactory performance of the aforementioned model seems to be an occasional observation. Therefore, this finding could not be regarded as indicative of the reliability and applicability of MSI as a means of evaluating the microbiological spoilage of oyster mushrooms.

Overall, the geometrical shape of the mushroom cap may constitute an important limitation for MSI. It is well known that the lighting reflectance is not uniformly distributed when the sample has spherical shape and this results to great spectral variability (Zhang et al, 2018). Similar limitations to the surface side, could be also arise for the gills due to the numerous physical cavities present, resulting in irregular reflectance. Gowen et al. (2008) investigated the application of hyperspectral imaging for detection of bruise damage on white mushrooms (*Agaricus bisporus*). In order to reduce the spectral variations caused by surface curvature, they applied and tested four spectral pre-processing methods to the mushroom spectra, including multiplicative scatter correction, maximum normalization, median normalization and mean normalization. Mean normalization constituted the best pre-treatment for decreasing spectral variability, but when the same pre-treatment was tested in the present data, the results were not promising (data not shown). Therefore, it could be overall opined that the phenomenon of microbiological spoilage in mushrooms is, most likely, difficult to be evaluated using spectral imaging technologies.

Table 6.4. Performance parameters of the partial least squares regression models, based on MSI data, for the total viable count prediction of fresh mushrooms. (*a*: slope, *b*: offset, r: correlation coefficient, RMSE: root mean square error (log CFU/g), R^2: coefficient of determination).

	Side	Duplicate samples	Test set	*a*	*b*	r	R^2	RMSE
MSI	Caps	No averaging	Dynamic	0.11	5.42	0.26	0.04	1.19
			Random	0.2	4.86	0.46	0.16	1.06
			Batch 1	0.29	4.37	0.51	0.23	0.98
			Batch 2	0.53	2.65	0.57	0.06	0.92
			Batch 3	0.25	5.55	0.54	-0.51	1.61
		Averaging	Dynamic	0.15	5.15	0.3	0.04	1.09
			Batch 1	0.46	3.22	0.66	0.38	0.8
			Batch 2	0.8	1.09	0.65	0.04	0.81
			Batch 3	0.33	5.03	0.7	-0.47	1.44
	Gills	No averaging	Dynamic	0.25	4.49	0.43	0.16	1.14
			Random	0.39	3.78	0.58	0.31	0.95
			Batch 1	0.35	4.19	0.46	0.12	1.05
			Batch 2	0.46	3.14	0.58	0.18	0.86
			Batch 3	0.25	5.24	0.51	-0.16	1.41
		Averaging	Dynamic	0.23	4.79	0.42	0.13	1.04
			Batch 1	0.63	2.33	0.76	0.58	0.66
			Batch 2	0.66	1.87	0.68	0.25	0.72
			Batch 3	0.3	4.96	0.55	-0.22	1.31

6.4. Conclusions

Rapid and non-invasive analytical technologies based on spectroscopy were, for the first time, applied for the evaluation of the microbiological spoilage of fresh oyster mushrooms. The results of the present study indicated that the application of FTIR and MSI is not promising for the evaluation of microbiological spoilage associated with such products. The models for

both mushroom sides (i.e. caps and gills) exhibited poor prediction performances, regardless of the applied data partitioning and spectral variability reduction.

Regarding the extensive physical and biological variability characterizing these products, a large-scale sampling of fresh mushrooms may constitute a better and integrated approach for extending the current knowledge. The methodologies used in sample preparation and sensor application should also be revised, maintaining the rapid and easy-to-use characteristics of the applied spectroscopy-based technologies. Moreover, additional computational analysis including, among others, the comparative analysis of various machine learning algorithms and data fusion approaches, could be also valuable for the development of reliable and robust predictive models that will better explain the process of microbiological spoilage of oyster mushrooms.

Acknowledgments

This work has been supported by the project "PhasmaFOOD", funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 732541.

Chapter 7

Discussion and future perspectives

Chapter 7

7.1. General discussion

Along with the growing production and consumption of fresh-cut and ready-to-eat produce, important challenges concerning the quality of these products continue to exist (Chaudry et al., 2018). However, specific knowledge about the diversity of produce-associated microbial communities related to spoilage is still scarce (Leff & Fierer, 2013; Gorni et al., 2015). Therefore, a better and deeper insight into the microbial communities, and the way various environmental factors influence and shape them, is definitely required to develop strategies for quality control and assurance (Juste et al., 2008; Leff & Fierer, 2013; Cao et al., 2017). Given also their notably limited shelf-life, which is to an important extent due to microbial activity, the establishment of rapid and effective methods for microbiological spoilage assessment is a first priority action for the fresh-cut food industry (Wang et al., 2018; Giannoglou et al., 2020).

The research undertaken in the present thesis initially aimed at exploring the microbial communities associated with the spoilage of fresh-cut produce and the changes occurring in their composition during storage at different temperatures through a culture-independent NGS technology (Section A). Although culture-dependent methods have been widely used in food microbiology, it is well established that they are not able to provide a broad and fine characterization of food microbial communities (Ercolini et al., 2013; Zhou et al., 2015; Edet et al., 2017). Subsequently, the application of spectroscopy-based technologies as alternative tools for the microbiological spoilage assessment of various fresh-cut produce commodities was also evaluated (Section B). Compared to conventional microbiological methods, these technologies provide minimal sample preparation, non-destructive sampling and rapid data acquisition (Nychas et al., 2016). Thus, they could allow real-time and onsite assessment at food industries and potentially at food retail and service premises.

7.1.1. Section A: Discussion

Numerous NGS studies, based mainly on metagenetic approaches, have investigated the plant microbiota associated with edible or inedible plant parts (Knief et al., 2014; Abdelfattah et al. 2018; Angeli et al., 2019). However, there is limited understanding of the

diversity and composition of the microbial communities of fresh and even more of fresh-cut produce (Leff & Fierer, 2013). The microbiota of such products changes following harvest, processing, packaging and storage (Jackson et al., 2015, Soderqvist et al., 2017; Gu et al., 2018). However, the microbial communities present at the time of purchase and consumption are more relevant to study, given the fact that they represent the communities to which consumers are exposed. In Chapters 2 and 3, the microbial communities of RTE pineapple and leafy vegetables were characterized, while the influence of temperature on their composition also was investigated using a metagenetic sequencing approach. The fresh-cut produce commodities were stored at different isothermal and dynamic temperature conditions. An important innovation of the present thesis was that the bacterial communities were characterized through *gyrB* sequencing providing species identification. Instead, all the published studies for fresh-cut produce utilize the widely used 16S rRNA analysis and remain at the taxonomic level of phylum, group or genus (Cao et al., 2017).

The study described in Chapter 2 reports on findings regarding the evolution of fungal (as the predominant spoilage microorganisms) and bacterial communities associated with RTE pineapple during storage under different temperature conditions. The fungal characterization was performed through a metagenetic amplicon sequencing approach, based on the ITS2 region, whereas bacterial identification was achieved through *gyrB* amplicon sequencing. Initially, a small scale comparative analysis of the culture-independent (metagenetic approach) and culture-dependent (plates) characterization of pineapple's fungal community indicated that a whole phylum (i.e. Basidiomycota), as wells as abundant Ascomycota species, were hardly or unsuccessfully detected in plates. Therefore, in accordance with literature, the applied culture-dependent method was evaluated as extremely biased in its ability to capture the fungal diversity of pineapple (Ercolini et al., 2013; Cao et al., 2017; Edet et al., 2017). With regard to the metagenetic analysis, the temperature and storage time showed a statistically significant impact in species richness which decreased over storage time and when the temperature reached higher levels. However, when the influence of different storage temperatures and times on fungal diversity and composition was investigated, a strong batch effect was revealed. Indeed, pineapple samples from two of the four studied batches (i.e. P3 and P4) had a quite similar fungal community (dominated by *Candida argentea* or *Hanseniaspora uvarum*), while samples from the other two batches (i.e. P1 and P2) displayed a set of completely different

fungal communities (dominated mostly by the species *Fusarium circinatum* or *Candida intermedia*, and in some cases by *Pichia fermentans* and *Meyerozyma caribbica*). These compositional variations may be attributed to the natural variability of plant-origin products, as well as the variability driven by various biological, pre-harvest and post- harvest factors (Rico et al., 2007; Leff & Fierer, 2013). A deeper analysis at batch level revealed that the impact of temperature and storage time varied on recovered fungal species, depending on their initial prevalence. Specifically, the abundant *F. circinatum* (phytopathogen) was gradually suppressed during storage, whereas other yeast species (*Pichia fermentans* or *C.argentea* at 4 ^{o}C, and *C. intermedia* at 8 and 12 ^{o}C) finally prevailed according to the temperature, indicating a potentially competitive action. The bioprotective activity of these yeasts against phytopathogenic molds, has been previously reported in the literature (Giobbe et al. 2007; Rosa-Magri et al. 2011). There were also species (e.g., *C. argentea*) which were initially dominant and remained as such throughout storage. As far as the subdominant bacterial microbiota is concerned, the *Pseudomonas* spp. was the most abundant genus, but bacteria of the Enterobacteriaceae family and lactic acid bacteria (LAB) also were present. However, the temperature and the batch factors did not significantly influence these microbial communities.

In Chapter 3, the bacterial epiphytic and endophytic communities associated with the spoilage of RTE rocket and baby spinach during storage at compliant and abusive temperatures were characterized using *gyrB* amplicon sequencing. Among the NGS studies concerning fresh or fresh-cut leafy vegetables, only few of them analysed both epiphytic and endophytic communities (Jackson et al., 2013, Soderqvist et al., 2017). The findings of the study indicated that *Pseudomonas* species were the most prevalent bacteria in both vegetables. The species diversity of baby spinach was significantly higher compared to that of rocket in terms of evenness. Furthermore, a striking difference in the bacterial community composition of the two vegetables was observed. Specifically, *Pseudomonas viridiflava* was highly abundant and occassionaly dominant in RTE rocket, while a new *Pseudomonas* species, namely *P. RIT357*, showed high abundances and even dominance in baby spinach. Moreover, *Pseudomonas fluorescens* and/or *Pseudomonas fragi* as well as *Janthinobacterium psychrotolerans*, and *Janthinobacterium lividum* were also highly abundant in some cases of baby spinach compared to rocket. The differences in species composition of the two vegetables may be attributed to variations in metabolites, physical characteristics, and symbiotic interactions with the host

(Redford et al., 2010; Leff & Fierer, 2013). However, the different batches within each type of vegetable differed slightly or largely in bacterial community composition, even in samples provided by the same supplier. The results also highlighted that the impact of temperature was significantly stronger in baby spinach evenness, while the storage time influenced more the rocket evenness. Similarly to the pineapple case, the impact of the aforementioned two environmental factors on species composition was not stronger than the batch impact. At batch level, no specific and clear impact of temperature and time was revealed on baby spinach communities. On the other hand, the impact of storage time was present in some batches of rocket. Interestingly, *Pseudomonas* species (*P. viridiflava* and *Pseudomonas syringae*), which are considered plant pathogens, became less abundant during storage, while members of the LAB group showed increased abundances. Indeed, LAB are also described in the literature as important bioprotective bacteria depending on the strain and the food product (Pothakos et al., 2014; Saraoui et al., 2016).

7.1.2. Section A: Conclusions and future perspectives

In the present section, the microbiota related to the fresh-cut produce spoilage during storage at different temperatures was successfully characterized at the species level, through metagenetics based on ITS and *gyrB* genetic markers for fungi and bacteria, respectively. A great variability in microbial communities among the different batches of the fresh-cut produce was observed, indicating that these commodities are very complex and unpredictable ecological niches. As it was previously mentioned, various factors including geographical origin, season of harvest, farming practices, as well as fresh-cut processing and storage conditions can shape distinctively the microbial communities and result in high compositional variations (Rastogi et al., 2012; Jackson et al., 2013; Leff & Fierer, 2013; Dees et al., 2015; Cao et al., 2017; Darlison et al., 2019; Tatsika et al., 2019). As a result of the extensive variability, the batch effect on microbial composition was stronger than that of temperature and storage time. Subsequently, a batch-level analysis was conducted with the aim to reveal any potential impact of storage temperature and/or time. In the case of pineapple, the impact of these two factors varied, according to the initial prevalent fungal species. Contrarily, the in-depth analysis did not reveal a specific and clear effect of temperature and time on baby spinach communities, while for rocket only the impact of storage time was observed in some cases.

Given all the above, a large-scale sampling of RTE produce commodities from various production facilities should allow for a broader and integrated analysis, resulting in better characterization of the diverse microbial communities specifically associated to the spoilage of these products. Further research is also necessary in order to unravel how important environmental factors may drive the spoilage of end products, before purchace and consumption. However, the successful distinction of the various batches of the studied fresh-cut produce commodities through the present metagenetic analysis could constitute an important tool for the industry. Specifically, this technology could provide a molecular fingerprint for each produce batch that would further contribute to the development of an advanced treacability system for quality management.

The results of the present thesis also reported potential microbial interactions during spoilage. The reduction of dominant species abundances were followed by the dominance of others. In pineapple, various yeasts finally dominated over a widely known pathogenic mold, which was initially prevalent. In the case of leafy vegetables, the potential suppression of bacterial species, often reported as plant pathogens, by LAB also was indicated. According to the scientific literature, several bacterial and yeast species or strains have been reported as natural microbial antagonists against fungal and bacterial pathogens of plants, but also human pathogens on a variety of harvested commodities (Giobbe et al. 2007; Sharma et al., 2009; Lopez-Velasco et al. 2010; Rosa-Magri et al., 2011; Soderqvist et al., 2017). However, the application and efficiency of biocontrol strategies has yet to be demonstrated in fresh-cut products. Indeed, spoilage is a complex biological phenomenon and little is known about the actual roles played by the microbial species involved in this procedure (Galimberti et al., 2015; Poirier et al., 2018). Even if valuable information can be gained through taxonomic description, conclusions are limited due to the lack of functional information. Microorganisms exhibit a wide genetic variability even within species, especially with respect to their metabolic pathways and host-interactive capabilities (Massart et al., 2015). Although, the metagenetic approach used in the present thesis is a powerful, cost-effective and easy to apply tool for taxonomic description, it does not provide any information on the genes present in the microbial communities (Massart et al., 2015). On the other hand, metagenomics can provide taxonomic and functional information by identifying the genes present in the food microbiome. Such functional analysis could yield a more accurate description and a better understanding of

the role of microbial species and their interactions in fresh-cut produce (Angeli et al. 2019). Although such analysis is more complex and costly in terms of sequencing and bioinformatics, it should be prioritised for a better understanding of the microbial ecology (Soderqvist et al., 2017). However, the integration of metagenomics and metatranscriptomics data is also important since the latter identify not only the highly abundant genes but also the most transcribed genes and pathways, further refining the functional understanding of microbiota's roles and functions. To conclude, the in-depth study of microbial communities and the various taxa (but mainly species) interactions with the aim to control and promote the growth of potential bioprotective/biocontrol agents is also important for the development of effective strategies concerning quality of RTE produce.

7.1.3. Section B: Discussion

Although spectroscopy-based technologies have been implemented in a broad range of food products (mainly meat products) with the aim to assess various aspects of quality, limited are the corresponding studies that have attempted to evaluate the microbiological spoilage in fresh produce, let alone fresh-cut and ready-to-eat (RTE) commodities (Suthiluk et al., 2008; Di Egidio et al., 2009). In Chapters 4, 5 and 6, different spectroscopy-based technologies were used for the assessment of microbiological quality of RTE pineapple, RTE leafy vegetables and fresh-cut oyster mushrooms, respectively. Given that microbial communities and their populations related to food spoilage depend on the product itself, as well as the applied storage conditions, the experiments were designed so as to include storage under different temperatures and, wherever applicable, under different packaging conditions (i.e. passive and active modified atmosphere packaging). Non-isothermal storage was also applied, since it is considered more realistic with regard to the conditions encountered in the cold chain for refrigerated food products (Ndraha et al., 2018). The products under study were subjected to parallel spectroscopy and microbiological analysis, aiming for the correlation of the data derived from the two methods and the development of appropriate predictive models. However, advanced computational tools are needed for the analysis of the collected data and the interpretation of results derived from these methods (Jollife & Cadima, 2016; Truong et al., 2019; Zhou et al., 2019). Therefore, various machine learning approaches were performed

and often comparatively evaluated in order to achieve the best possible and realistic model performance for the assessment of microbiological spoilage of fresh-cut produce.

In Chapter 4, various analytical technologies (sensors), namely Fourier-transform infrared (FTIR), near-infrared (NIR), fluorescence (FLUO) and visible (VIS) spectroscopies as well as multispectral imaging (MSI), were utilized in evaluating the microbiological (total viable count, TVC), but also sensory (colour, odour, texture) quality of RTE pineapple (*Ananas comosus*). The suitability of different machine learning algorithms (linear and non-linear) and sensors for monitoring the various features was assessed comparing the prediction performances of the models developed. The potentials of The Unscrambler software and the online machine learning ranking platform, SorfML, also were explored. Since various easy-to-use software are commercially provided, it is important to learn the potentials and the limitations of each platform, realizing that conflicting outputs are often possible (Nunes et al., 2015). For the two statistical software, different data pre-processing, data partitioning schemes and cross-validation methods were applied for model development. Besides all these distinct approaches, the results indicated similar trends about the ability of sensors and algorithms to assess the quality of RTE pineapple. Specifically, for TVC prediction, almost all the models developed with the partial least squares regression (PLSR) algorithm showed relatively satisfactory performances in both The Unscrambler and SorfML software. However, the MSI, and mainly FLUO and VIS model exhibited better performances in The Unscrambler, with R^2 values varying from 0.54 to 0.58 and RMSE of prediction not above 0.53 log CFU/g. In SorfML, FLUO and PLSR combination also exhibited the best performance with slight differences compared to the majority of the PLSR models.Apart from PLSR, all the support vector machine (SVM) Linear models were also satisfactory for TVC prediction, exhibiting RMSE and R^2 values close and occasionally (in the case of VIS data) improved to those of PLSR models in SorfML. On the other hand, non-linear algorithms exhibited poor model performances. The applicability of NIR spectroscopy was assessed using only The Unscrambler software, and the corresponding model exhibited the poorest performance of all the PLSR models developed in both software. Consequently, the results indicated that the assessment of the microbiological spoilage of RTE pineapple could be potentially achieved by the majority of the studied spectroscopy-based technologies. In the case of sensory quality, the conduction of a more integrated and representative analysis of sensory attributes is needed for

solid conclusions to be drawn. However, FLUO and MSI seem to be also promising for the evaluation of pineapple's odour.

A comparative analysis of non-invasive sensors and machine learning approaches for evaluating the microbiological spoilage of RTE leafy vegetables was conducted and presented in Chapter 5. Sensors based on FTIR, NIR and VIS spectroscopies as well as MSI, were used for the assessment of TVC and *Pseudomonas* spp. populations of RTE rocket and baby spinach salads stored in passive MAP. Along with the various sensors, two machine learning algorithms, PLSR and support vector regression (SVR, radial), and two distinct data partitioning approaches [i.e. random and dynamic (as test set) data partitioning] were utilized. The results demonstrated that the random data partitioning scheme resulted in better performances in the case of rocket and in almost all cases of sensor/algorithm combinations. On the other hand, in the case of baby spinach, the random partitioning approach resulted often in better performances for MSI and NIR, but in similar performances when FTIR and VIS were applied. With regard to the different algorithms used, the predictive power was similar for the two algorithms in baby spinach, with few exceptions. Contrarily, in rocket, the SVR algorithm provided considerably or slightly (depending on the data partitioning approach) better models for the FTIR, VIS and NIR sensors but lower performance for the MSI sensor. Overall, the microbiological spoilage of baby spinach was better assessed by models derived mainly from the VIS (i.e. slope: 0.74; offset: 2.04; R^2: 0.65; RMSE: 0.4) sensor, while FTIR (i.e. slope: 0.71; offset: 2.14; R^2: 0.70; RMSE: 0.72) spectroscopy and MSI (with similar or even better performances in some cases) were found to be more suitable in the case of rocket. In Chapter 5, FTIR spectroscopy and MSI were applied also for the evaluation of the microbiological spoilage of RTE baby spinach stored under active MAP conditions. In contrast to the above results in baby spinach, FTIR spectroscopy and MSI appeared to be promising tools, since the performance indices indicated good performances, which were also the highest performances among all the developed models. Finally, the results indicated that a tailor-made sample and data analysis workflow is needed for the two leafy vegetables. Therefore, it is evident that the different types of vegetables, as well as the datasets derived from the same product but corresponding to different storage conditions, should be treated distinctively in terms of sensor and computational analysis application.

Concerning Chapter 6, the microbiological spoilage of oyster mushrooms (*Pleurotus ostreatus*) was assessed using FTIR spectroscopy and MSI. Prior to PLSR, various data partitioning methods were tested. All the developed PLSR models for both mushroom sides (cap and gills) and both sensors exhibited quite poor prediction performances, regardless of the applied data partitioning schemes, feature selection and reduction of spectral variability. The results derived from the applied data analysis indicated that FTIR spectroscopy and MSI are not promising tools for the evaluation of the microbiological spoilage of oyster myshrooms. The poor applicability of the two sensors could be attributed to a combination of potential methodological failures, limitations of analytical technologies and structural difficulties of the product resulting, in turn, in the observed extensive spectral variability.

7.1.4. Section B: Conclusions and future perspectives

The majority of the tested spectroscopy-based technologies coupled with machine learning showed promising results for the evaluation of the microbiological spoilage of fresh-cut produce, with the exception of oyster mushrooms. Currently, these technologies could provide additional information to the well-established microbiological and molecular analyses, allowing for the time-efficient, non-destructive and onsite screening of fresh-cut produce samples in the production line. However, future research is necessary for these methods to be eligible for the accurate and precise evaluation of fresh-cut produce spoilage and to be able to replace the above conventional retrospective and expensive methods.

Interestingly, all the fresh-cut produce datasets showed a great microbiological and spectral variability among batches, as well as between the two biological replicates (duplicate samples originated from different packages). This excessive variability is very common in plant-origin products due to the strong impact of biological characteristics (e.g., cultivar and geographical region), as well as various pre-harvest (agricultural practices) and post-harvest (processing and storage conditions) factors on the final product attributes and quality (Martınez-Sanchez et al., 2006; Conte et al., 2008; Medina et al., 2012; Garrido et al., 2015; Zhang et al., 2018). Even among produce samples from the same field, strong variations may be present due the great natural variability of such products (Rico et al., 2007). Therefore, the administration of large-scale experiments is advised to improve the characterization of the extensive variability of fresh produce commodities. Samples derived from different cultivars,

geographical regions and seasons, as well as from different producers and fresh-cut processing facilities (with different pre-processing and packaging steps) should allow for a better representation and integration of the encountered variability in model calibration (Zhang et al., 2018). Moreover, the development of open databases that could be continuously enriched with data derived from various certified laboratories seems to also constitute a great option for developing more robust and universal models.

During the model development procedure, various data analysis approaches were evaluated along with the tested spectroscopy-based sensors in the present thesis in order to achieve the best possible model performances. First of all, the application of a data partitioning scheme based on data derived from dynamic temperature conditions was proposed. The data derived from the dynamic conditions include the information from all temperatures and depict realistic temperature fluctuations during produce distribution in the food supply, and, hence should allow for the development of meaningful predictive models (Tsakanikas et al., 2018). Indeed, the results demonstrated that when model external validation was performed based on the dynamic temperature data, the models performance was in some cases similar to the one attained when validation was performed on samples randomly selected. Moreover, different machine learning algorithms were comparatively evaluated for their suitability to assess the microbiological spoilage of produce commodities. Among the machine learning algorithms tested, PLSR exhibited relatively satisfactory or even good predictive power and in some cases the best model performances. PLSR is a linear method widely used for spectroscopy datasets that encounter dimensionality issues, strong collinearity and noise (Wold et al., 2001; Mehmood et al., 2012). However, the present results indicated that other algorithms, even non-linear, were also suitable for the evaluation of microbiological spoilage. Specifically, SVM Linear (for regression and classification) utilized in pineapple and SVR (non-linear) based on radial basis function kernel in leafy vegetables exhibited similar or even better performances compared to PLSR. Therefore, additional computational analysis including, among others, the utilization of various data pre-processing methods, and selection of the most informative (in each case) wavelengths and data fusion approaches, should also be applied and comparatively evaluated, allowing for the development of more reliable and robust predictive models. Interestingly, the applicability of the various sensors and machine learning approaches in the evaluation of microbiological spoilage seems to be food-specific. According to Chapter 5,

although the two leafy vegetables tested seem quite similar to each other, different sensors and algorithms were more appropriate for each vegetable type. The results were also differentiated among datasets derived from the same food product but different packaging conditions. Therefore, it is evident that the different commodities, as well as same products subjected to different storage conditions, should be treated distinctively in terms of sensor and machine learning approach application.

Overall, the utilization of spectroscopy-based technologies in food science applications requires the involvement of different scientific disciplines including food microbiology, chemistry, physics, engineering and statistics. A multi-disciplinary approach is anticipated to contribute to a better understanding of the complex microbiological spoilage mechanisms related to fresh-cut produce, and, thereby, to a focused and optimized application of spectroscopy-based technologies in food quality control. The conduction of informative metabolomic analysis, focused on the identification and quantification of specific metabolites related to food spoilage, would provide valuable information for the biochemical activity of spoilage microorganisms, and would support future research with regard to non-invasive analytical technologies such as the ones studied in the present thesis.

References

Abdelfattah, A., Wisniewski, M., Giulia, M., Destri, L., 2016. Metagenomic Analysis of Fungal Diversity on Strawberry Plants and the Effect of Management Practices on the Fungal Community Structure of Aerial Organs. PLoS One 1–17.
https://doi.org/10.1371/journal.pone.0160470

Abdelfattah, A., Malacrinò, A., Wisniewski, M., Cacciola, S.O., Schena, L., 2018. Metabarcoding : A powerful tool to investigate microbial communities and shape future plant protection strategies. Biol. Control 120, 1–10.
https://doi.org/10.1016/j.biocontrol.2017.07.009

Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D.,Cech, M., Chilton, J., Clements, D., Coraor, N., Gruning, B.A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., Blankenberg, D., 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses : 2018 update. Nucleic Acids Res 46, 537–544.
https://doi.org/10.1093/nar/gky379

Al-Dosary, N.M.N., Al-Hamed, S.A., Aboukarima, A.M., 2019. K-nearest neighbors method for prediction of fuel consumption in tractor – chisel plow systems. Eng. Agricola 39 (6), 729–736. http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v39n6p729-736/2019

Al-holy, M.A., Lin, M., Alhaj, O.A., Abu-Goush, M.H., 2015. Discrimination between Bacillus and Alicyclobacillus Isolates in Apple Juice by Fourier Transform Infrared Spectroscopy and Multivariate Analysis. J. Food Sci. 00, 1–6.
https://doi.org/10.1111/1750-3841.12768

Al-Jowder, O., Defernez, M., Kemsley, E.K., Wilson, R.H., 1999. Mid-infrared spectroscopy and chemometrics for the authentication of meat products. J. Agric. Food Chem. 47, 3210–3218. https://doi.org/10.1021/jf981196d

Allende, A., Luo, Y., Mcevoy, J.L., Artés, F., Wang, C.Y., 2004. Microbial and quality changes in minimally processed baby spinach leaves stored under super atmospheric oxygen and modified atmosphere conditions. Postharvest Biol. Technol. 33, 51–59. https://doi.org/10.1016/j.postharvbio.2004.03.003

Ammor, M.S., Argyri, A., Nychas, G.J.E., 2009. Rapid monitoring of the spoilage of minced beef stored under conventionally and active packaging conditions using Fourier transform infrared spectroscopy in tandem with chemometrics. Meat Sci. 81, 507–514. https://doi.org/10.1016/j.meatsci.2008.10.015

Angeli, D., Sare, A.R., Jijakli, M.H., Pertot, I., Massart, S., 2019. Insights gained from metagenomic shotgun sequencing of apple fruit epiphytic microbiota. Postharvest Biol. Technol. 153, 96–106. https://doi.org/10.1016/j.postharvbio.2019.03.020

Argyri, A.A., Jarvis, R.M., Wedge, D., Xu, Y., Panagou, E.Z., Goodacre, R., Nychas, G.E., 2013. A comparison of Raman and FT-IR spectroscopy for the prediction of meat spoilage. Food Control 29, 461-470. https://doi.org/10.1016/j.foodcont.2012.05.040

Artes, F., Allende, A., 2014. Minimal Processing of Fresh Fruit, Vegetables, and Juices, in: Sun D.W (Ed.), Emerging Technologies for Food Processing. Academic Press, USA, pp. 583-597.

Auer, L., Mariadassou, M., O'Donohue, M., Klopp, C., Hernandez-Raquet, G., 2017. Analysis of large 16S rRNA Illumina datasets: Impact of singleton read filtering on microbial community description. Mol. Ecol. Resour. 17(6):e122±e32. https://doi.org/10.1111/1755-0998.12700

Awad, M., Khanna, R., 2015. Efficient learning machines: theories, concepts, and applications for engineers and system designers. New York: Apress Open.

Baselice, A., Colantuoni, F., Lass, D.A., Nardone, G., Stasi, A., 2017. Trends in EU consumers' attitude towards fresh-cut fruit and vegetables. Food Qual. Prefer. 59, 87–96. https://doi.org/10.1016/j.foodqual.2017.01.008

Barbin, D.F., Kaminishikawahara, C.M., Soares, A.L., Mizubuti, I.Y., Grespan, M., Shimokomaki, M., Hirooka, E.Y., 2015. Prediction of chicken quality attributes by near infrared spectroscopy. Food Chem. 168, 554–560. https://doi.org/10.1016/j.foodchem.2014.07.101

Barker, M., Rayens, W., 2003. Partial least squares for discrimination. J. Chemom. 17, 166–173. https://doi.org/10.1002/cem.785

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. 43 (5), 772–777.

Bekiaris, G., Tagkouli, D., Koutrotsios, G., Kalogeropoulos, N., Zervakis, G.I., 2020. *Pleurotus* Mushrooms Content in Glucans and Ergosterol Assessed by ATR-FTIR Spectroscopy and Multivariate Analysis. Foods *9*, 535. https://doi.org/10.3390/foods9040535

Belgiu, M., Dragut, L., 2016. Random forest in remote sensing : A review of applications and future directions. SPRS J Photogramm Remote Sens 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bophela, K.N., Petersen, Y., Carolee. T. Bull, C.T., Teresa. A. Coutinho T.A., 2020. Identification of Pseudomonas isolates associated with bacterial canker of stone fruit trees in the Western Cape , South Africa 104 (3), 1–48. Plant Disease 104 (3), 882-892. https://doi.org/10.1094/PDIS-05-19-1102-RE

Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. Training algorithm for optimal margin classifiers. Proc. Fifth Annu. ACM Work. Comput. Learn. Theory 144–152. https://doi.org/10.1145/130385.130401

Brandily, M.L., Monbet, V., Bureau, B., Boussard-Plédel, C., Loréal, O., Adam, J.L., Sire, O., 2011. Identification of foodborne pathogens within food matrices by IR spectroscopy. Sensors Actuators, B Chem. 160, 202–206. https://doi.org/10.1016/j.snb.2011.07.034

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Bureau, S., Cozzolino, D., Clark, C.J., 2019. Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables : A review. Postharvest Biol. Technol. 148, 1–14. https://doi.org/10.1016/j.postharvbio.2018.10.00, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST + : architecture and applications. BMC Bioinform. 9, 1–9. https://doi.org/10.1186/1471-2105-10-421

Caldera, L., Franzetti, L., Van Coillie, E., De Vos, P., Stragier, P., De Block, J., & Heyndrickx, M., 2016. Identification, enzymatic spoilage characterization and proteolytic activity quantification of *Pseudomonas* spp. isolated from different foods. Food Microbiol., 54, 142-153.

Camps, C., Christen, D., 2009. Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. LWT - Food Sci. Technol. 42, 1125–1131. https://doi.org/10.1016/j.lwt.2009.01.015

Cao, Y., Fanning, S., Proos, S., Jordan, K., Srikumar, S., 2017. A Review on the Applications of Next Generation Sequencing Technologies as Applied to Food-Related Microbiome Studies. Front. Microbiol. 8, 1–16. https://doi.org/10.3389/fmicb.2017.01829

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., Mcdonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. correspondence QIIME allows analysis of high- throughput community sequencing. Nat. Methods 7 (5), 335-336.
https://doi.org/10.1038/nmeth0510-335

Carstensen, J.M., Hansen, J.F., 2003. An apparatus and a method of recording an image of an object. Patent family EP1051660.

Castro-Ibanez, I., Gil, M.I., Allende, A., 2017. Ready-to-eat vegetables: Current problems and potential solutions to reduce microbial risk in the production chain. LWT - Food Sci. Technol. 85, 284-292. http://dx.doi.org/10.1016/j.lwt.2016.11.073

Cocolin, L., Mataragas, M., Bourdichon, F., Doulgeraki, A., Pilet, M.-F., Jagadeesan, B., Rantsiou, K., Phister, T., 2018. Next generation microbiological risk assessment meta-omics: The next need for integration. Int. J. Food Microbiol. 287, 10–17.

Cauchie, E., Delhalle, L., Taminiau, B., Tahiri, A., Korsak, N., Burteau, S., Fall, P.A., Farnir, F., Baré, G., 2020. Assessment of Spoilage Bacterial Communities in Food Wrap and Modified Atmospheres-Packed Minced Pork Meat Samples by 16S rDNA Metagenetic Analysis. Front. Microbiol. 10, 1–17. https://doi.org/10.3389/fmicb.2019.03074

Chanprasartsuk, O., Prakitchaiwattana, C., Sanguandeekul, R., Fleet, G.H., 2010. Autochthonous yeasts associated with mature pineapple fruits, freshly crushed juice and their ferments ; and the chemical changes during natural fermentation. Bioresour. Technol. 101, 7500–7509. https://doi.org/10.1016/j.biortech.2010.04.047

Chaudhry, M.M.A., Amodio, M.L., Babellahi, F., De Chiara, M.L.V., Amigo Rubio, J.M., Colelli, G., 2018. Hyperspectral imaging and multivariate accelerated shelf-life testing

(MASLT) approach for determining shelf-life of rocket leaves. J. Food Eng. 238, 122–133. https://doi.org/10.1016/j.jfoodeng.2018.06.017

Coldea, T.E., Socaciu, C., Fetea, F., Ranga, F., Pop, R.M., Florea, M., 2013. Rapid quantitative analysis of ethanol and prediction of methanol content in traditional fruit brandies from romania, using FTIR spectroscopy and chemometrics. Not. Bot. Horti Agrobot. Cluj-Napoca 41, 143–149. https://doi.org/10.15835/nbha4119000

Conte, A., Conversa, G., Scrocco, C., Brescia, I., Laverse, J., Elia, A., Del Nobile, M. A., 2008. Influence of growing periods on the quality of baby spinach leaves at harvest and during storage as minimally processed produce. Postharvest Biol. Technol., 50(2-3), 190-196. https://doi.org/10.1016/j.postharvbio.2008.04.003

Cooley, M.B., Chao, D., Mandrell, R.E., 2006. Escherichia coli O157 : H7 Survival and Growth on Lettuce Is Altered by the Presence of Epiphytic Bacteria. J. Food Prot. 69, 2329–2335.

Cortes, C., Vapnik, V., 1995. Support-Vector Networks. Mach. Learn. 20, 273–297.

Cozzolino, D., 2015. Foodomics and infrared spectroscopy : from compounds to functionality. Curr. Opin. Food Sci. 4, 39-43. https://doi.org/10.1016/j.cofs.2015.05.003

da Cruz Almeida, E.T., de Medeiros Barbosa, I., Tavares, J.F., Barbosa-Filho, J.M., Magnani, M., and de Souza, E.L., 2018. Inactivation of Spoilage Yeasts by *Mentha spicata* L. and *M.* × *villosa* Huds. Essential Oils in Cashew, Guava, Mango, and Pineapple Juices. Front. Microbiol. 9, 1–12. https://doi.org/10.3389/fmicb.2018.01111

Darlison, J., Mogren, L., Rosberg, A. K., Grudén, M., Minet, A., Liné, C., Mieli, M., Bengtssona, T., Håkansson, A., Uhlig, E., Becherc, P.G., Karlsson, M. & Alsanius B.W., 2019. Leaf mineral content govern microbial community structure in the phyllosphere of spinach (Spinacia oleracea) and rocket (Diplotaxis tenuifolia). *Sci. Total Environ., 675*, 501-512. https://doi.org/10.1016/j.scitotenv.2019.04.254

Davis, R., Irudayaraj, J., Reuhs, B.L., Mauer, L.J., 2010. Detection of E. coli O157:H7 from ground beef using fourier transform infrared (FTIR) spectroscopy and chemometrics. J. Food Sci. 75, 340–346. https://doi.org/10.1111/j.1750-3841.2010.01686.x

Dees, M.W., Lysøe, E., Nordskog, B., Brurberg, M.B., 2015. Bacterial communities associated with surfaces of leafy greens: Shift in composition and decrease in richness over time. Appl. Environ. Microbiol. 81, 1530–1539. https://doi.org/10.1128/AEM.03470-14

de la Mata, P., Dominguez-Vidal, A., Bosque-Sendra, J.M., Ruiz-Medina, A., Cuadros-Rodríguez, L., Ayora-Cañada, M.J., 2012. Olive oil assessment in edible oil blends by means of ATR-FTIR and chemometrics. Food Control 23, 449–455. https://doi.org/10.1016/j.foodcont.2011.08.013

De Sousa Marques, A., Nicácio, J.T.N., Cidral, T.A., de Melo, M.C.N., de Lima, K.M.G., 2013. The use of near infrared spectroscopy and multivariate techniques to differentiate Escherichia coli and Salmonella Enteritidis inoculated into pulp juice. J. Microbiol. Methods 93, 90–94. https://doi.org/10.1016/j.mimet.2013.02.003

den Besten, H.M.W., Amézquita, A., Bover-Cid, S., Dagnas, S., Ellouze, M., Guillou, S., Nychas, G., O'Mahony, C., Pérez-Rodriguez, F., Membré, J.M., 2018. Next generation of microbiological risk assessment: Potential of omics data for exposure assessment. Int. J. Food Microbiol. 287, 18–27. https://doi.org/10.1016/j.ijfoodmicro.2017.10.006

Di Cagno, R., Cardinali, G., Minervini, G., Antonielli, L., Giuseppe-Rizzello, C., Ricciuti, P., Gobbetti, M., 2010. Taxonomic structure of the yeasts and lactic acid bacteria microbiota of pineapple (*Ananas comosus* L. Merr.) and use of autochthonous starters for minimally processing. Food Microbiol. 27, 381–389. https://doi.org/10.1016/j.fm.2009.11.012

Di Egidio, V., Sinelli, N., Limbo, S., Torri, L., Franzetti, L., Casiraghi, E., 2009. Evaluation of shelf-life of fresh-cut pineapple using FT-NIR and FT-IR spectroscopy. Postharvest Biol. Technol. 54, 87–92. https://doi.org/10.1016/j.postharvbio.2009.06.006

Ding, T., Rahman, S.M.E., Oh, D., 2011. Inhibitory effects of low concentration electrolyzed water and other sanitizers against foodborne pathogens on oyster mushroom. Food Control 22, 318–322. https://doi.org/10.1016/j.foodcont.2010.07.030

Dixit, Y., Casado-Gavalda, M.P., Cama-Moncunill, R., Cama-Moncunill, X., Markiewicz-Keszycka, M., Cullen, P.J., Sullivan, C., 2017. Developments and Challenges in Online NIR Spectroscopy for Meat Processing. Compr. Rev. Food Sci. Food Saf. 16, 1172–1187. https://doi.org/10.1111/1541-4337.12295

Doona, C. J., Feeherry, F. E., Feng, H., Grove, S., Krishnamurthy, K., Lee, A., & Kustin, K., 2015. Combining sanitizers and non-thermal processing technologies to improve fresh-cut produce safety. In Electron Beam Pasteurization and Complementary Food Processing Technologies (pp.95-125). Elsevier Inc. https://doi.org/10.1533/9781782421085.2.95

Dos Santos Souza, C.R., de Oliveira Barbosa, A.C., Ferreira, C.F., Souza, F.V.D., de Souza Rocha, L., de Souza, E.H., de Oliveira, S.A.S., 2019. Diversity of microorganisms associated to *Ananas* spp. from natural environment, cultivated and *ex situ* conservation areas. Sci. Hortic. 243, 544–551. https://doi.org/10.1016/j.scienta.2018.09.015

Duan, C., Chen, C., Khan, M.N., Liu, Y., Zhang, R., Lin, H., Cao, L., 2014. Non-destructive determination of the total bacteria in flounder fillet by portable near infrared spectrometer. Food Control 42, 18–22. https://doi.org/10.1016/j.foodcont.2014.01.023

Edet, U.O., Antai, S.P., Brooks, A.A., Asitok, A.D, Enya, O., Japhet, F.H, 2017. An Overview of Cultural, Molecular and Metagenomic Techniques in Description of Microbial Diversity. JAMB 7(2): 1-19. https://doi.org/10.9734/JAMB/2017/37951

Efenberger-Szmechtyk, M., Nowak, A., Kregiel, D., 2018. Implementation of chemometrics in quality evaluation of food and beverages. Crit. Rev. Food Sci. Nutr. 58, 1747–1766. https://doi.org/10.1080/10408398.2016.1276883

EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards), Koutsoumanis, K., Alvarez-Ordonez, A., Bolton, D., Bover-Cid, S., Chemaly, M., Davies, R., De Cesare, A., Herman, L., Hilbert, F., Lindqvist, R., Nauta, M., Peixe, L., Ru, G., Simmons, M., Skandamis, P., Suffredini, E., Jordan, K., Sampers, I., Wagner, M., Da Silva Felicio, M.T., Georgiadis, M., Messens, W., Mosbach-Schulz, O., Allende, A., 2020. Scientific Opinion on the public health risk posed by *Listeria monocytogenes* in frozen fruit and vegetables including herbs, blanched during processing. EFSA J 18(4):6092, 102 pp. https://doi.org/10.2903/j.efsa.2020.6092

Ellis, D.I., Goodacre, R., 2001. Rapid and quantitative detection of the microbial spoilage of muscle foods: current status and future trends. Trends Food Sci Technol 12, (11) 414-424. https://doi.org/10.1016/S0924-2244(02)00019-5

Ellis, D.I., Broadhurst, D., Kell, D.B., Jem, J., Goodacre, R., Rowland, J.J., 2002. Rapid and Quantitative Detection of the Microbial Spoilage of Meat by Fourier Transform Infrared Spectroscopy and Machine Learning. Appl. Environ. Microbiol. 68, 2822–2828. https://doi.org/10.1128/AEM.68.6.2822

Ercolini, D., 2013. High-Throughput Sequencing and Metagenomics: Moving Forward in the Culture-Independent Analysis of Food Microbial Ecology. Appl. Environ. Microbiol. 79, 3148–3155. https://doi.org/10.1128/aem.00256-13

Ergun, M., 2006. Fresh-cut Physiology and Factors Contributing to the Quality of Fresh-cut Produce. KSU. J. Eng. Sci. 9(2), 164-169.

Erkmen, O., Bozoglu, T.F., 2016. Spoilage of Vegetables and Fruits. In Food Microbiology: Principles into Practice, 1st ed. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119237860.ch20

Escudie, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., Maman, S., Hernandez-Raquet, G., Combes, S. and Pascal, G., 2018. Sequence analysis FROGS: Find, Rapidly, OTUs with Galaxy Solution. Bioinformatics 34, 1287–1294. https://doi.org/10.1093/bioinformatics/btx791

Estelles-Lopez, L., Ropodi, A., Pavlidis, D., Fotopoulou, J., Gkousari, C., Peyrodie, A., Panagou, E., Nychas, G.J., Mohareb, F., 2017. An automated ranking platform for machine learning regression models for meat spoilage prediction using multi-spectral imaging and metabolic profiling. Food Res. Int. 99, 206–215. https://doi.org/10.1016/j.foodres.2017.05.013

Esquerre, C., Gowen, A.A., Downey, G., O'Donnell, C.P., 2012. Wavelength selection for development of a near infrared imaging system for early detection of bruise damage in mushrooms (Agaricus bisporus). J. Near Infrared Spectrosc. 20, 537–546. doi: 10.1255/jnirs.1014

FDA (U.S. Department of Health and Human Services Food and Drug Administration Center for Food Safety and Applied Nutrition), 2018. Guide to Minimize Food Safety Hazards of Fresh-cut Produce: Draft Guidance for Industry. FDA-2018-D-3583.

Fengou, L.-C., Spyrelli, E., Lianou, A., Tsakanikas, P., Panagou, E.Z., Nychas, G.-J.E., 2019a. Estimation of Minced Pork Microbiological Spoilage through Fourier Transform Infrared and Visible Spectroscopy and Multispectral Vision Technology. Foods 8, 238. https://doi.org/10.3390/foods8070238

Fengou, L.C., Lianou, A., Tsakanikas, P., Gkana, E.N., Panagou, E.Z., Nychas, G.J.E., 2019b. Evaluation of Fourier transform infrared spectroscopy and multispectral imaging as means of estimating the microbiological spoilage of farmed sea bream. Food Microbiol. 79, 27–34. https://doi.org/10.1016/j.fm.2018.10.020

Francis, G.A/, Gallone, A., Nychas, G.J., Sofos, J.N., Colelli, G., Amodio, M.L., Spano, G., 2012. Factors affecting quality and safety of fresh-cut produce. Crit Rev Food Sci Nutr. 52(7), 595-610. doi: 10.1080/10408398.2010.503685.

Franzetti, L. and Scarpellini, M., 2007. Characterisation of *Pseudomonas* spp. isolated from foods. Ann. Microbiol. 57, 39–47. https://doi.org/10.1007/BF03175048

Gaglio, R., Miceli, A., Sardina, M.T., Francesca, N., Moschetti, G., Settanni, L., 2019. Evaluation of microbiological and physico - chemical parameters of retail ready - to - eat mono - varietal salads. J Food Process Preserv. 1–12. https://doi.org/10.1111/jfpp.13955

Garrido, Y., Tudela, J.A., Gil, M.I., 2015. Time of day for harvest and delay before processing affect the quality of minimally processed baby spinach. Postharvest Biol. Technol. 110, 9–17. http://dx.doi.org/10.1016/j.postharvbio.2015.07.002

Garrido, Y., Tudela, J.A., Hernández, J.A., Gil, M.I., 2016. Modified atmosphere generated during storage under light conditions is the main factor responsible for the quality changes of baby spinach. Postharvest Biol. Technol. 114, 45–53.

Gaston, E., Frías, J.M., Cullen, P.J., O'Donnell, C.P., Gowen, A.A., 2010. Prediction of Polyphenol Oxidase Activity Using Visible Near-Infrared Hyperspectral Imaging on Mushroom (Agaricus bisporus) Caps. J. Agric. Food Chem. 58 (10), 6226-6233.
DOI: 10.1021/jf100501q

Gautam, R., Vanga, S., Ariese, F., Umapathy, S., 2015. Review of multidimensional data processing approaches for Raman and infrared spectroscopy. EPJ Tech. Instrum.
https://doi.org/10.1140/epjti/s40485-015-0018-6

Giannoglou, M., Stergiou, P., Dimitrakellis, P., Gogolides, E., Stoforos, N.G., Katsaros, G., 2020. Effect of Cold Atmospheric Plasma processing on quality and shelf-life of ready-to-eat rocket leafy salad. Innov. Food Sci. Emerg. Technol. 66, 102502.

https://doi.org/10.1016/j.ifset.2020.102502

Giobbe, S., Marceddu, S., Scherm, B., Zara, G., Mazzarello, V.L., Budroni, M., Migheli, Q., 2007. The strange case of a biofilmm-forming strain of *Pichia fermentans*, which controls *Monilinia* brown rot on apple but is pathogenic on peach fruit. FEMS Yeast Res 7 1389–1398. https://doi.org/10.1111/j.1567-1364.2007.00301.x

Goodacre, R., Broadhurst, D., Smilde, A.K., Kristal, B.S., Baker, J.D., Beger, R., Bessant, C., Connor, S., Capuani, G., Craig, A., Ebbels, T., Kell, D.B., Manetti, C., Newton, .J., Paternostro, G., Somorjai, R., Sjostrom, M., Trygg, J., Wulfert, F., 2007. Proposed minimum reporting standards for data analysis in metabolomics. Metabolomics 3, 231–241. https://doi.org/10.1007/s11306-007-0081-3

Gowen, A.A., Donnell, C.P.O., Taghizadeh, M., Cullen, P.J., Frias, J.M., Downey, J., 2008. Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (Agaricus bisporus). J. Chemometrics 22, 259–267. https://doi.org/10.1002/cem.1127

Gram, L., Ravn, L., Rasch, M., Bruhn, J.B., Christensen, A.B., Givskov, M., 2002. Food spoilage--interactions between food spoilage bacteria. Int. J. Food Microbiol. 78, 79–97.

Granato, D., de Araújo Calado, V.Ô.M., Jarvis, B., 2014. Observations on the use of statistical methods in Food Science and Technology. Food Res. Int. 55, 137–149. https://doi.org/10.1016/j.foodres.2013.10.024

Granato, D., Putnik, P., Kovačević, D.B., Santos, J.S., Calado, V., Rocha, R.S., Cruz, A.G. Da, Jarvis, B., Rodionova, O.Y., Pomerantsev, A., 2018. Trends in Chemometrics: Food Authentication, Microbiology, and Effects of Processing. Compr. Rev. Food Sci. Food Saf. 17, 663–677. https://doi.org/10.1111/1541-4337.12341

Grewal, M.K., Jaiswal, P., Jha, S.N., 2015. Detection of poultry meat specific bacteria using FTIR spectroscopy and chemometrics. J. Food Sci. Technol. 52, 3859–3869. https://doi.org/10.1007/s13197-014-1457-9

Gromski, P.S., Muhamadali, H., Ellis, D.I., Xu, Y., Correa, E., Turner, M.L., Goodacre, R., 2015. A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. Anal. Chim. Acta 879, 10–23. https://doi.org/10.1016/j.aca.2015.02.012

Gong, X., Skrivergaard, S., Korsgaard, B.S., Schreiber, L., Marshall, I.P.G., Finster, K., Schramm, A., 2017. High quality draft genome sequence of *Janthinobacterium psychrotolerans* sp . nov ., isolated from a frozen freshwater pond. Stand. Genomic Sci. 1–7. https://doi.org/10.1186/s40793-017-0230-x

Goodwin, S., Mcpherson, J.D., Mccombie, W.R., 2016. Coming of age : ten years of next-generation sequencing technologies. Nature 17, 333-351.https://doi.org/10.1038/nrg.2016.49

Gorni, C., Allemand, D., Rossi, D., Mariani, P., 2015. Microbiome profiling in fresh-cut products. Trends Food Sci. Technol. 46, 295–301. https://doi.org/10.1016/j.tifs.2015.10.013

Gu, G., Ottesen, A., Bolten, S., Ramachandran, P., Reed, E., Rideout, S., Luo, Y., Patel, J., Brown, E., Nou, X., 2018. Shifts in spinach microbial communities after chlorine washing and storage at compliant and abusive temperatures. Food Microbiol. *73*, 73-84. https://doi.org/10.1016/j.fm.2018.01.002

Guo, Q., Wu, W., Massart, D.L., 1999. The robust normal variate transform for pattern recognition with near-infrared data. Anal. Chim. Acta 382 (1), 87–103. https://doi.org/10.1016/S0003-2670(98)00737-5

Guzmán, E., Baeten, V., Pierna, J.A.F., García-Mesa, J.A., 2015. Determination of the olive maturity index of intact fruits using image analysis. J. Food Sci. Technol. 52, 1462–1470. https://doi.org/10.1007/s13197-013-1123-7

Handschur, M., Pinar, G., Gallist, B., Lubitz, W., Haslberger, A.G., 2005. Culture free DGGE and cloning based monitoring of changes in bacterial communities of salad due to processing. Food Chem. Toxicol. 43, 1595–1605. https://doi.org/10.1016/j.fct.2005.05.009

Hassoun, A., Sahar, A., Lakhal, L., Aït-kaddour, A., 2019. Fluorescence spectroscopy as a rapid and non-destructive method for monitoring quality and authenticity of fish and meat products : Impact of different preservation conditions. LWT - Food Sci. Technol. 103, 279–292. https://doi.org/10.1016/j.lwt.2019.01.021

He, H.-J., Wu, D., Da-Wen Sun, D.-W., 2013. Non-destructive and rapid analysis of moisture distribution in farmed Atlantic salmon (Salmo salar) fillets using visible and near-infrared hyperspectral imaging. Innov Food Sci Emerg Technol 18, 237–245. http://dx.doi.org/10.1016/j.ifset.2013.02.009

He, H.J., Sun, D.W., 2015a. Toward enhancement in prediction of Pseudomonas counts distribution in salmon fillets using NIR hyperspectral imaging. LWT - Food Sci. Technol. 62, 11-18. https://doi.org/10.1016/j.lwt.2015.01.036

He, H.J., Sun, D.W., 2015b. Inspection of harmful microbial contamination occurred in edible salmon flesh using imaging technology. J. Food Eng. 150, 82–89. https://doi.org/10.1016/j.jfoodeng.2014.10.012

Hibbing, M.E., Fuqua, C., Parsek, M.R., Peterson, S.B., 2010. Bacterial competition: surviving and thriving in the microbial jungle. Nat Rev Microbiol 8, 15–25. https://doi.org/10.1038/nrmicro2259

Hoffman, C.S., Winston, F., 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene 57, 267–272.

Hu, L., Yin, C., Ma, S., Liu, Z., 2018. Rapid detection of three quality parameters and classification of wine based on Vis-NIR spectroscopy with wavelength selection by ACO and CARS algorithms. Spectrochim Acta A Mol Biomol Spectrosc 205, 574-581. https://doi.org/10.1016/j.saa.2018.07.054.

Huyan, Z., Ding, S., Liu, X., Yu, X., 2018. Authentication and adulteration detection of peanut oils of three flavor types using synchronous fluorescence spectroscopy. Anal. Methods 10, 3207–3214. https://doi.org/10.1039/c8ay00837j

Hunter, P.J., Hand, P., Pink, D., Whipps, J.M., Bending, G.D., 2010. Both Leaf Properties and Microbe-Microbe Interactions Influence Within-Species Variation in Bacterial Population Diversity and Structure in the Lettuce (*Lactuca* Species) Phyllosphere. Appl. Environ. Microbiol. 76, 8117–8125. https://doi.org/10.1128/AEM.01321-10

Ibrahim, N.F., Hawa, M., Nik, M., Izham, M., Nor, M., Zakaria, L., 2017. Characterization of *Fusarium* spp. associated with pineapple fruit rot and leaf spot in Peninsular Malaysia. J. Phytopathol. 718–726. https://doi.org/10.1111/jph.12611

Jackson, C.R., Randolph, K.C., Osborn, S.L., Tyler, H.L., 2013. Culture dependent and independent analysis of bacterial communities associated with commercial salad leaf vegetables. BMC Microbiol. 13, 1–12. https://doi.org/10.1186/1471-2180-13-274

Jackson, C.R., Stone, B.W.G., Tyler, H.L., 2015. Emerging Perspectives on the Natural Microbiome of Fresh Produce Vegetables. Agriculture 5, 170–187. https://doi.org/10.3390/agriculture5020170

Jacobs, A., Schalk, P., Wyk, V.A.N., Marasas, W.F.O., Wingfield, B.D., Wingfield, M.J., Coutinho, T.A., 2010. *Fusarium ananatum* sp. nov. in the *Gibberella fujikuroi* species complex

from pineapples with fruit rot in South Africa. Fungal Biol. 114, 515–527. https://doi.org/10.1016/j.funbio.2010.03.013

Jacquot, S., Karoui, R., Abbas, K., Lebecque, A., Bord, C., Aït-Kaddour, A., 2015. Potential of Multispectral Imager to Characterize Anisotropic French PDO Cheeses: A Feasibility Study. Int. J. Food Prop. 18, 213–230. https://doi.org/10.1080/10942912.2013.828746

James, J.B., Ngarmsak, T., 2010. Processing of fresh-cut tropical fruitsand vegetables: A technical guide, in: Rolle, R.S. (Ed.). Food and Agriculture Organization of the United Nations, Regional Office for Asia and the Pacific, Bangkok, 2010. RAP PUBLICATION 2010/16

Jayawardena, B., & Silva, H. A. P., 2014. Effect of Cinnamic Acid and Modified Atmosphere Packaging on Storage Life of *Pleurotus Ostreatus* (Oyster Mushrooms). JOSUK, *8*, 19–32. http://doi.org/10.4038/josuk.v8i0.5230

Jideani, A.I.O., Anyasi, T.A.M., Mchau, G.R.A., Udoro, E.O., Onipe, O.O., 2017. Processing and Preservation of Fresh-Cut Fruit and Vegetable Products, in: Kahramanoglu, I., (Ed.), Postharvest Handling, IntechOpen, DOI: 10.5772/intechopen.69763. Available from: https://www.intechopen.com/books/postharvest-handling/processing-and-preservation-of-fresh-cut-fruit-and-vegetable-products.

Jollife, I.T., Cadima, J., 2016. Principal component analysis: A review and recent developments. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374. https://doi.org/10.1098/rsta.2015.0202

Juste, A., Thomma, B.P.H.J., Lievens. B., 2008. Recent advances in molecular techniques to study microbial communities in food-associated matrices and processes. Food Microbiol. 25, 745–761. https://doi.org/10.1016/j.fm.2008.04.009

Kader, A.A., 2002. Postharvest Biology and Technology: An Overview. Postharvest Technology of Horticultural Crops, third ed. University of California Agriculture and Natural Resources, 3311.

Kiala, Z., Odindi, J., Mutanga, O., Peerbhay, K., 2016. Comparison of partial least squares and support vector regressions for predicting leaf area index on a tropical grassland using hyperspectral data. J. Appl. Remote Sens. 10(3), 036015. https://doi.org/10.1117/1.JRS.10.036015

Kergourlay, G., Taminiau, B., Daube, G., Champomier Vergès, M.C., 2015. Metagenomic insights into the dynamics of microbial communities in food. Int. J. Food Microbiol. 213, 31–39. https://doi.org/10.1016/j.ijfoodmicro.2015.09.010

Klerks, M.M., Franz, E., Gent-Pelzer, M. Van, Zijlstra, C., Bruggen, A.HC. Van, 2007. Differential interaction of Salmonella enterica serovars with lettuce cultivars and plant-microbe factors influencing the colonization efficiency. ISME 620–631. https://doi.org/10.1038/ismej.2007.82

Knief, C., 2014. Analysis of plant microbe interactions in the era of next generation sequencing technologies. Front. Plant Sci. 5, 1–23. https://doi.org/10.3389/fpls.2014.00216

Koukounaras, A., Bantis, F., Karatolos, N., Melissas, C., Vezyroglou, A., 2020. Influence of Pre-Harvest Factors on Postharvest Quality of Fresh-Cut and Baby Leafy Vegetables. Agronomy 10, 172. https://doi.org/10.3390/agronomy10020172

Krivoshiev, G.P., Chalucova, R.P., Moukarev, M.I., 2000. A Possibility for Elimination of the Interference from the Peel in Nondestructive Determination of the Internal Quality of Fruit and Vegetables by VIS / NIR Spectroscopy. Lebensm.-Wiss. u.-Technol. 33, 344–353. doi:10.1006/fstl.2000.0669

Kumar, N., Bansal, A., Sarma, G.S., Rawal, R.K., 2014. Chemometrics tools used in analytical chemistry: An overview. Talanta 123, 186–199. https://doi.org/10.1016/j.talanta.2014.02.003

Kumar, S., Krishnani, K.K., Bhushan, B., Brahmane, M.P., 2015. Metagenomics : Retrospect and Prospects in High Throughput Age 2015. Biotechnol. Res. Int. 2015. http://dx.doi.org/10.1155/2015/121735

Largeteau, M.L., & Savoie, J.-M., 2010. Microbially induced diseases of Agaricus bisporus: biochemical mechanisms and impact on commercial mushroom production. Appl Microbiol Biotechnol, 86, 63-73.

Leff, J.W., Fierer, N., 2013. Bacterial Communities Associated with the Surfaces of Fresh Fruits and Vegetables. PLoS One 8, 1–9. https://doi.org/10.1371/journal.pone.0059310

Leneveu-Jenvrin, C., Quentin, B., Assemat, S., Hoarau, M., Meile, J., Remize, F., 2020. Changes of Quality of Minimally-Processed Pineapple (*Ananas comosus*, var. 'Queen Victoria') during Cold Storage : Fungi in the Leading Role. Microorganisms 8, 185. https://doi.org//microorganisms8020185

Leo, A., Tow, W.W., Hill, J., English, J.R., Ford, R., 2013. Novel Salmonella spp. Diagnostic Markers Based on the gyrB Gene and its Application on Food and Environmentally-Derived Water Samples. J. Adv. Biotechnol. Bioeng. 1, 30–39.

Li, B., Lecourt, J., Bishop, G., 2018. Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction—a review. Plants 7, 1–20. https://doi.org/10.3390/plants7010003

Lianou, A., Mencattini, A., Catini, A., Di Natale, C., Nychas, G.J.E, Martinelli, E., Panagou, E.Z., 2019. Online Feature Selection for Robust Classification of the Microbiological Quality of Traditional Vanilla Cream by Means of Multispectral Imaging. Sensors, 19(19):4071. doi: 10.3390/s19194071.

Lima, G.B.L., Rosa, C.A., Johann, S., Vieira, M.L.A., Gomes, F.C.O., 2019. Yeasts isolated from tropical fruit ice creams : diversity, antifungal susceptibility and adherence to buccal epithelial cells. Braz. J. Food Technol., Campinas, v. 22, e2018197. https://doi.org/10.1590/1981-6723.19718 1–10

Lin, M., Rasco, B.A., Cavinato, A.G., Al-Holy, M., 2009. Infrared (IR) Spectroscopy—Near-Infrared Spectroscopy and Mid-Infrared Spectroscopy, in: Sun, D.W. (Ed.), Infrared Spectroscopy for Food Quality Analysis and Control, Academic Press, pp.119-143. https://doi.org/10.1016/B978-0-12-374136-3.00006-7.

Liu, C., Liu, W., Chen, W., Yang, J., Zheng, L., 2015. Feasibility in multispectral imaging for predicting the content of bioactive compounds in intact tomato fruit. Food Chem. 173, 482–488. https://doi.org/10.1016/j.foodchem.2014.10.052

Lohumi, S., Lee, S., Lee, H., Cho, B.-K., 2015. A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. Trends Food Sci. Technol. 46, 85–98. https://doi.org/10.1016/j.tifs.2015.08.003

Løkke, M.M., Seefeldt, H.F., Skov, T., Edelenbos, M., 2013. Color and textural quality of packaged wild rocket measured by multispectral imaging. Postharvest Biol. Technol. 75, 86–95. https://doi.org/10.1016/j.postharvbio.2012.06.018

Lopez-Velasco, G., Davis, M., Boyer, R.R., Williams, R.C., Ponder, M.A., 2010. Alterations of the phylloepiphytic bacterial community associated with interactions of Escherichia coli O157 : H7 during storage of packaged spinach at refrigeration temperatures. Food Microbiol. 27, 476–486. https://doi.org/10.1016/j.fm.2009.12.010

Lopez-Velasco, G., Welbaum, G.E., Boyer, R.R., Mane, S.P., Ponder, M.A., 2011. Changes in spinach phylloepiphytic bacteria communities following minimal processing and refrigerated storage described using pyrosequencing of 16S rRNA amplicons. J. Appl. Microbiol. 110, 1203–1214. https://doi.org/10.1111/j.1365-2672.2011.04969.x

Lunadei, L., Diezma, B., Lleó, L., Ruiz-Garcia, L., Cantalapiedra, S., Ruiz-Altisent, M., 2012. Monitoring of fresh-cut spinach leaves through a multispectral vision system. Postharvest Biol. Technol. 63, 74–84. https://doi.org/10.1016/j.postharvbio.2011.08.004

Lunadei, L., Galleguillos, P., Diezma, B., Lleó, L., Ruiz-Garcia, L., 2011. A multispectral vision system to evaluate enzymatic browning in fresh-cut apple slices. Postharvest Biol. Technol. 60, 225–234. https://doi.org/10.1016/j.postharvbio.2011.02.001

Ma, L., Zhang, M., Bhandari, B., Gao, Z., 2017. Recent developments in novel shelf-life extension technologies of fresh-cut fruits and vegetables. Trends Food Sci. Technol. 64, 23–38. https://doi.org/10.1016/j.tifs.2017.03.005

Maciel, N.O.P., Piló, F.B., Freitas, L.F.D., Gomes, F.C.O., Johann, S., Nardi, R.M.D., Lachance, M., Rosa, C.A., 2013. The diversity and antifungal susceptibility of the yeasts isolated from coconut water and reconstituted fruit juices in Brazil. Int. J. Microbiol. 160, 201–205. https://doi.org/10.1016/j.ijfoodmicro.2012.10.012

Mahe, F., Rognes, T., Quince, C., de Vargas, C., Dunthorn, M., 2015. Swarm v2 : highly-scalable and high-resolution amplicon clustering. PeerJ 3:e1420. https://doi.org/10.7717/peerj.1420

Manthou, E., Lago, S.L., Dagres, E., Lianou, A., Tsakanikas, P., Panagou, E.Z., Anastasiadi, M., Mohareb, F., Nychas, G.J.E., 2020. Application of spectroscopic and multispectral imaging technologies on the assessment of ready-to-eat pineapple quality : A performance evaluation study of machine learning models generated from two commercial data analytics tools. Comput. Electron. Agric. 175, 105529. https://doi.org/10.1016/j.compag.2020.105529

Manthou, E., Coeuret, G., Chaillou, S., Nychas, G. J. E., 2021. Evolution of fungal community associated with ready-to-eat pineapple during storage under different temperature conditions. Food Microbiol. 97, 103736. https://doi.org/10.1016/j.fm.2021.103736

Martınez-Sanchez, A., Allende, A., Bennett, R.N., Ferreres, F., Gil, M.I., 2006. Microbial, nutritional and sensory quality of rocket leaves as affected by different sanitizers. Postharvest Biol. Technol. 42, 86–97. https://doi.org/10.1016/j.postharvbio.2006.05.010

Massart, S., Martinez-Medina, M., Jijakli M.H., 2015. Biological control in the microbiome era: Challenges and opportunities. Biol. Control 89, 98–108. http://dx.doi.org/10.1016/j.biocontrol.2015.06.003

McMurdie, P.J., Holmes, S., 2013. phyloseq : An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One 8. https://doi.org/10.1371/journal.pone.0061217

Medina, M.S., Tudela, J.A., Marín, A., Allende, A., Gil, M.I., 2012. Short postharvest storage under low relative humidity improves quality and shelf-life of minimally processed baby spinach (Spinacia oleracea L.). Postharvest Biol. Technol. 67, 1–9. https://doi.org/10.1016/j.postharvbio.2011.12.002

Mehmood, T., Liland, K.H., Snipen, L., Sæbø, S., 2012. A review of variable selection methods in Partial Least Squares Regression. Chemom. Intell. Lab. Syst. 118, 62–69. https://doi.org/10.1016/j.chemolab.2012.07.010

Meenu, M., Xu, B., 2019. Application of vibrational spectroscopy for classification, authentication and quality analysis of mushroom: A concise review. Food Chem. 289, 545–557. https://doi.org/10.1016/j.foodchem.2019.03.091

Montero-Calderón, M., Rojas-Graü, M.A., Martín-Belloso, O., 2008. Effect of packaging conditions on quality and shelf-life of fresh-cut pineapple (*Ananas comosus*). Postharvest Biol. Technol. 50, 182–189. https://doi.org/10.1016/j.postharvbio.2008.03.014

Morris, C.E., Kinkel, L.L., Xiao, K., Prior, P., Sands, D.C., 2007. Surprising niche for the plant pathogen *Pseudomonas syringae*. Infect Genet Evol. 7(1):84-92.
doi: 10.1016/j.meegid.2006.05.002.

Nastou, A., Rhoades, J., Smirniotis, P., Makri, I., Kontominas, M., Likotra, E., 2012. Efficacy of household washing treatments for the control of Listeria monocytogenes on salad vegetables. Int. J. Food Microbiol. 159, 247–253.
https://doi.org/10.1016/j.ijfoodmicro.2012.09.003

Ndraha, N., Hsiao, H.-I., Vlajic, J., Yang, M.-F., Lin, H.-T. V., 2018. Time-temperature abuse in the food cold chain: Review of issues, challenges, and recommendations, Food Control 89, 12-21. https://doi.org/10.1016/j.foodcont.2018.01.027

Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy : A review. Postharvest Biol. Technol. 46, 99–118.
https://doi.org/10.1016/j.postharvbio.2007.06.024

Nicolaou, N., Goodacre, R., 2008. Rapid and quantitative detection of the microbial spoilage in milk using Fourier transform infrared spectroscopy and chemometrics. Analyst, 133, 1424–1431. doi: 10.1039/b804439b.

Nilsson, R.H., Larsson, K.-H., Taylor, A.F.S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., Kennedy, P., Picard, K., Glockner, F.O., Tedersoo, L., Saar, I., Koljalg, U., Abarenkov, K., 2019. The UNITE database for molecular identification of fungi : handling dark taxa and parallel taxonomic classifications. Nucleic Acids Res 47, 259–264.
https://doi.org/10.1093/nar/gky1022

Nunes, C.A., Alvarenga, V.O., de Souza Sant'Ana, A., Santos, J.S., Granato, D., 2015. The use of statistical software in food science and technology: Advantages, limitations and misuses. Food Res. Int. 75, 270–280. https://doi.org/10.1016/j.foodres.2015.06.011

Nychas, G.J.E., Skandamis, P.N., Tassou, C.C., Koutsoumanis, K.P., 2008. Meat spoilage during distribution. Meat Sci. 78, 77–89. https://doi.org/10.1016/j.meatsci.2007.06.020

Nychas, G.J.E., Panagou, E.Z., Mohareb, F., 2016. Novel approaches for food safety management and communication. Curr. Opin. Food Sci. 12, 13–20. https://doi.org/10.1016/j.cofs.2016.06.005

Oliveira, M., Abadias, M., Usall, J., Torres, R., Teixidó, N., Viñas, I., 2015. Application of modified atmosphere packaging as a safety approach to fresh-cut fruits and vegetables-A review. Trends Food Sci. Technol. 46, 13–26. https://doi.org/10.1016/j.tifs.2015.07.017

Oto, N., Oshita, S., Makino, Y., Kawagoe, Y., Sugiyama, J., Yoshimura, M., 2013. Non-destructive evaluation of ATP content and plate count on pork meat surface by fluorescence spectroscopy. Meat Sci. 93, 579–585. https://doi.org/10.1016/j.meatsci.2012.11.010

Padrón-Mederos, M., Rodríguez-Galdón, B., Díaz-Romero, C., Lobo-Rodrigo, M.G., Rodríguez-Rodríguez, E.M., 2020. Quality evaluation of minimally fresh-cut processed pineapples. LWT - Food Sci. Technol. 129, 109607. https://doi.org/10.1016/j.lwt.2020.109607

Panagou, E.Z., Papadopoulou, O., Carstensen, J.M., Nychas, G.J.E., 2014. Potential of multispectral imaging technology for rapid and non-destructive determination of the microbiological quality of beef filets during aerobic storage. Int. J. Food Microbiol. 174, 1–11. https://doi.org/10.1016/j.ijfoodmicro.2013.12.026

Pantanella, F., Berlutti, F., Passariello, C., Sarli, S., Morea, C., Schippa, S., 2007. Violacein and biofilm production in *Janthinobacterium lividum*. J Appl Microbiol. 102(4):992-9. doi: 10.1111/j.1365-2672.2006.03155.x.

Papadopoulou, O., Panagou, E.Z., Tassou, C.C., Nychas, G.J.E., 2011. Contribution of Fourier transform infrared (FTIR) spectroscopy data on the quantitative determination of minced pork meat spoilage. Food Res. Int. 44, 3264–3271. https://doi.org/10.1016/j.foodres.2011.09.012

Papaianni, M., Paris, D., Woo, S. L., Fulgione, A., Rigano, M. M., Parrilli, E., Tutino, M. L., Marra, R., Manganiello, G., Casillo, A., Limone, A., Zoina, A., Motta, A., Lorito, M., Capparelli, R., 2020. Plant Dynamic Metabolic Response to Bacteriophage Treatment After *Xanthomonas campestris* pv. *campestris* Infection. Frontiers in microbiology, 11, 732. https://doi.org/10.3389/fmicb.2020.00732

Pradas-Baena, J.M. Moreno-Rojas, M.D. Luque de Castro, 2015. Effect of Processing on Active Compounds in Fresh-Cut Vegetables, in Preedy V. (Ed.), Processing and Impact on Active Components in Food. Academic Press, pp. 3-10. https://doi.org/10.1016/B978-0-12-404699-3.00001-9.

Pilone, V., Stasi, A., Baselice, A., 2017. Quality preferences and pricing of fresh-cut salads in Italy: new evidence from market data. Br. Food J. 119, 1473–1486. https://doi.org/10.1108/BFJ-09-2016-0419

Poirier, S., Rue, O., Peguilhan, R., Coeuret, G., Zagorec, M., Champomier-Vergès, M.-C., 2018. Deciphering intra-species bacterial diversity of meat and seafood spoilage microbiota using gyrB amplicon sequencing: A comparative analysis with 16S rDNA V3-V4 amplicon sequencing. PLoS One 13(9): e0204629. https://doi.org/10.1371/journal.pone.0204629

Poirier, S., Luong, N.-D. M., Anthoine, V., Guillou, S., Membré, J.-M., Moriceau, N., Rezé, S., Zagorec, M., Feurer, C., Frémaux, B., Jeuge, S., Robieu, E., Champomier-Vergès, M., Coeuret, G., Cauchie, E., Daube, G., Korsak, N., Coroller, L., Desriac, N., Desmonts, M.-H., Gohier, R., Werner, D., Loux, V., Rué, O., Dohollou, M.-H., Defosse, T., Chaillou, S., 2020. Large-scale multivariate dataset on the characterization of microbiota diversity, microbial growth dynamics, metabolic spoilage volatilome and sensorial profiles of two industrially

produced meat products subjected to changes in lactate concentration and packaging atmosphere. Data Brief, 30, 105453. https://doi.org/10.1016/j.dib.2020.105453

Pothakos, V., Nyambi, C., Zhang, B.Y., Papastergiadis, A., De Meulenaer, B., Devlieghere, F., 2014. Spoilage potential of psychrotrophic lactic acid bacteria (LAB) species: Leuconostoc gelidum subsp. gasicomitatum and Lactococcus piscium, on sweet bell pepper (SBP) simulation medium under different gas compositions. Int J Food Microbiol. 178:120-9. doi: 10.1016/j.ijfoodmicro.2014.03.012.

Power, A. C., Chapman, J., Chandra, S., and Cozzolino, D., 2019. Ultraviolet-visible spectroscopy for food quality analysis, in: Zhong, J., Wang, X. (Eds.), Evaluation technologies for food quality. Woodhead Publishing, Duxford, United Kingdom, pp. 91-104. https://doi.org/10.1016/B978-0-12-814217-2.00006-8

Prieto, N., Ross, D.W., Navajas, E.A., Nute, G.R., Richardson, R.I., Hyslop, J.J., Simm, G., Roehe, R., 2009. On-line application of visible and near infrared reflectance spectroscopy to predict chemical-physical and sensory characteristics of beef quality. Meat Sci. 83, 96–103. https://doi.org/10.1016/j.meatsci.2009.04.005

Qadri, O.S., Yousuf, B., Srivastava, A.K., 2015. Fresh-cut fruits and vegetables: Critical factors influencing microbiology and novel approaches to prevent microbial risks-A review. Cogent Food Agric. 1, 1–11. https://doi.org/10.1080/23311932.2015.1121606

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Glo, F.O., Yarza, P., Peplies, J., Glockner, F.O., 2013. The SILVA ribosomal RNA gene database project : improved data processing and web-based tools. Nucleic Acids Res 41, 590–596. https://doi.org/10.1093/nar/gks1219

Ragaert, P., Devlieghere, F., Debevere, J., 2007. Role of microbiological and physiological spoilage mechanisms during storage of minimally processed vegetables. Postharvest Biol. Technol. 44, 185–194. https://doi.org/10.1016/j.postharvbio.2007.01.001

Rastogi, G., Sbodio, A., Tech, J.J., Suslow, T. V., Coaker, G.L., Leveau, J.H.J., 2012. Leaf microbiota in an agroecosystem: Spatiotemporal variation in bacterial community composition on field-grown lettuce. ISME J. 6, 1812–1822. https://doi.org/10.1038/ismej.2012.32

Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y., Fierer, N., 2010. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. Environmental microbiology, 12(11), 2885–2893. https://doi.org/10.1111/j.1462-2920.2010.02258.x

Rico, D., Martın-Diana, A.B., Barat, J.M., Barry-Ryan, C., 2007. Extending and measuring the quality of fresh-cut fruit and vegetables: a review. Trends Food Sci Technol 18, 373-386.
Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH : a versatile open source tool for metagenomics. PeerJ 4:e2584. https://doi.org/10.7717/peerj.2584

Rinnan, A., Van den Berg, F., Engelsen, S.E., 2009. Review of the most common pre-processing techniques for near-infrared spectra. Trends Analyt Chem 28 (10), 1201-1222.

Ripolles-Avila, C., Martínez-garcia, M., Capellas, M., Yuste, J., Fung, D.Y.C., Rodríguez-Jerez, J.-J., 2020. From hazard analysis to risk control using rapid methods in microbiology : A practical approach for the food industry. Compr Rev Food Sci Food Saf. 19, 1877–1907. https://doi.org/10.1111/1541-4337.12592

Roberts, J.J., Cozzolino, D., 2016. An Overview on the Application of Chemometrics in Food Science and Technology-An Approach to Quantitative Data Analysis. Food Anal. Methods 9, 3258–3267. https://doi.org/10.1007/s12161-016-0574-7

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH : a versatile open source tool for metagenomics. PeerJ 4:e2584. https://doi.org/10.7717/peerj.2584

Rojas-Graü, M.A.,Garner, E., Martín-Belloso, O., 2010. The Fresh-Cut Fruit and Vegetables Industry: Current Situation and Market Trends, in: Martin-Belloso, O., Fortuny, R.S (Eds.), Advances in Fresh-Cut Fruits and Vegetables Processing. CRC Press, Boca Raton, pp. 1-11.

Ropodi, A.I., Panagou, E.Z., Nychas, G.J.E., 2017. Multispectral imaging (MSI): A promising method for the detection of minced beef adulteration with horsemeat. Food Control 73, 57–63. https://doi.org/10.1016/j.foodcont.2016.05.048

Ropodi, A.I., Panagou, E.Z., Nychas, G.J.E., 2016. Data mining derived from food analyses using non-invasive/non-destructive analytical techniques; determination of food authenticity, quality & safety in tandem with computer science disciplines. Trends Food Sci. Technol. 50, 11–25. https://doi.org/10.1016/j.tifs.2016.01.011

Ropodi, A.I., Pavlidis, D.E., Mohareb, F., Panagou, E.Z., Nychas, G.J.E., 2015. Multispectral image analysis approach to detect adulteration of beef and pork in raw meats. Food Res. Int. 67, 12–18. https://doi.org/10.1016/j.foodres.2014.10.032

Rosa-Magri, M.M., Tauk-Tornisielo, S.M., Ceccato-, S.R., 2011. Bioprospection of Yeasts as Biocontrol Agents Against Phytopathogenic Molds. Braz. Arch. Biol. Technol. 54, 1–5.

Rosberg, A.K., Darlison, J., Mogren, L., Alsanius, B.W., 2020. Commercial wash of leafy vegetables do not significantly decrease bacterial load but leads to shifts in bacterial species composition. Food Microbiol. 103667. https://doi.org/10.1016/j.fm.2020.103667

Rudi, K., Flateland, S.L., Hanssen, J.F., Bengtsson, G., Nissen, H., 2002. Development and Evaluation of a 16S Ribosomal DNA Array-Based Approach for Describing Complex Microbial Communities in Ready-To-Eat Vegetable Salads Packed in a Modified Atmosphere. Appl. Environ. Microbiol. 68, 1146–1156. https://doi.org/10.1128/AEM.68.3.1146

Sajben, E., Manczinger, L., Nagy, A., Kredics, L., & Vágvölgyi, C. (2011). Characterization of pseudomonads isolated from decaying sporocarps of oyster mushroom. Microbiol. Res., 166, 255-267.

Saminathan, T., García, M., Ghimire, B., Lopez, C., 2018. Metagenomic and Metatranscriptomic Analyses of Diverse Watermelon Cultivars Reveal the Role of Fruit Associated Microbiome in Carbohydrate Metabolism and Ripening of Mature Fruits. Front. Plant Sci. 9, 1–13. https://doi.org/10.3389/fpls.2018.00004

Sandhya, 2010. Modified atmosphere packaging of fresh produce: Current status and future needs. LWT - Food Sci. Technol. 43, 381–392. https://doi.org/10.1016/j.lwt.2009.05.018

Santana, C.C., Dantas Vanetti, M.C., & Megumi Kasuya, M.C., 2008. Microbial growth and colour of minimally processed shiitake mushroom stored at different temperatures. Int. J. Food Sci. Technol., 43, 1281-1285.

Saraiva, C., Vasconcelos, H., De Almeida, J.M.M.M., 2016. A chemometrics approach applied to fourier transform infrared spectroscopy (ftir) for monitoring the spoilage of fresh salmon (*Salmo salar*) stored under modified atmospheres. Int. J. Food Microbiol. https://doi.org/10.1016/j.ijfoodmicro.2016.10.038

Saraoui, T., Leroi, F., Björkroth, J., Pilet, M.F., 2016. Lactococcus piscium: a psychrotrophic lactic acid bacterium with bioprotective or spoilage activity in food-a review. J Appl Microbiol. 121(4):907-18. doi: 10.1111/jam.13179.

Saranraj, P., Stella, D., Reetha, D., 2012. Microbial spoilage of vegetables and its control measures: a review. Int. J. Nat. Prod. Sci, 2, 1-12.

Sarris, P.F., Trantas, E.A., Mpalantinaki, E., Ververidis, F., Goumas, D.E., 2012. Pseudomonas viridiflava, a Multi Host Plant Pathogen with Significant Genetic Variation at the Molecular Level. PLoS one, 7(4), e36090. https://doi.org/10.1371/journal.pone.0036090

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur : Open-Source , Platform-Independent , Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. 75 (23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schuenzel, K.M., Harrison, M.A., 2002. Microbial Antagonists of Foodborne Pathogens on Fresh, Minimally Processed Vegetables. J. Food Prot. 65, 1909–1915.

Sharma, R. R., Singh, D., & Singh, R., 2009. Biological control of postharvest diseases of fruits and vegetables by microbial antagonists: A review. Biol. Control, 50(3), 205-221. https://doi.org/10.1016/j.biocontrol.2009.05.001

Silverman, B.W., Jones, M.C., 1989. E. fix and jl hodges (1951): An important contribution to nonparametric discriminant analysis and density estimation: Commentary on fix and hodges (1951). Int. Stat. Rev. 57, 233-247.

Sikorska, E., Khmelinskii, I., Sikorski, M., 2019. Fluorescence spectroscopy and imaging instruments for food quality evaluation, in: Zhong, J., Xichang Wang, X. (Eds.), Evaluation Technologies for Food Quality. Woodhead Publishing Series in Food Science, Technology and Nutrition, Elsevier Inc., pp. 491-533. https://doi.org/10.1016/B978-0-12-814217-2.00019-6

Söderqvist, K., Osman, O.A., Wolff, C., Bertilsson, S., Vågsholm, I., Boqvist, S., 2017. Emerging microbiota during cold storage and temperature abuse of ready-to-eat salad. Infect. Ecol. Epidemiol. 7, 1328963. https://doi.org/10.1080/20008686.2017.1328963

Soliva-Fortuny, R.C., Martín-Belloso, O., 2003. New advances in extending the shelf-life of fresh-cut fruits: A review. Trends Food Sci. Technol. 14, 341–353. https://doi.org/10.1016/S0924-2244(03)00054-2

Spyrelli, E.D., Doulgeraki, A.I., Argyri, A.A., Tassou, C.C., Panagou, E.Z., Nychas, G.J.E., 2020. Implementation of Multispectral Imaging (MSI) for Microbiological Quality Assessment of Poultry Products. Microorganisms 1–14. https://doi.org/10.3390/microorganisms8040552

Sravan Kumar, S., Manoj, P., Giridhar, P., 2015. Fourier transform infrared spectroscopy (FTIR) analysis, chlorophyll content and antioxidant properties of native and defatted foliage of green leafy vegetables. J. Food Sci. Technol. 52, 8131–8139. https://doi.org/10.1007/s13197-015-1959-0

Subasinghe, R.M., Samarajeewa, A.D., Scroggins, R., Beaudette, L.A., 2019. Evaluation of denaturing gradient gel electrophoresis (DGGE) and next generation sequencing (NGS) in combination with enrichment culture techniques to identify bacteria in commercial microbial-based products. J. Microbiol. Methods 161, 118–130. https://doi.org/10.1016/j.mimet.2019.04.017

Subramanian, A., Alvarez, V.B., Harper, W.J., Rodriguez-Saona, L.E., 2011. Monitoring amino acids, organic acids, and ripening changes in Cheddar cheese using Fourier-transform infrared spectroscopy. Int. Dairy J. 21, 434–440. https://doi.org/10.1016/j.idairyj.2010.12.012

Suhandy, D., 2009. Nondestructive measurement of soluble solids content in pineapple fruit using short wavelength near infrared (SW-NIR) spectroscopy. Int. J. Appl. Eng. Res. 4, 107–114.

Suhandy, D., Yulia, M., 2017. The Use of Partial Least Square Regression and Spectral Data in UV-Visible Region for Quantification of Adulteration in Indonesian Palm Civet Coffee. Int. J. Food Sci. 2017. https://doi.org/10.1155/2017/6274178

Sun, Y., Li, W., 2017. Effects of different cling films on freshness of *Pleurotus ostreatus*. Food Measure 11, 592–597. https://doi.org/10.1007/s11694-016-9427-8

Suthiluk, P., Saranwong, S., Kawano, S., Numthuam, S., Satake, T., 2008. Original article Possibility of using near infrared spectroscopy for evaluation of bacterial contamination in shredded cabbage. Int. J. Food Sci. Technol. 43, 160–165. https://doi.org/10.1111/j.1365-2621.2006.01416.x

Taghizadeh, M., Gowen, A., O'Donnell, C.P., 2009. Prediction of white button mushroom (Agaricus bisporus) moisture content using hyperspectral imaging. Sens. & Instrumen. Food Qual. 3, 219–226. DOI 10.1007/s11694-009-9088-y

Taghizadeh, M., Gowen, A., Ward, P., O'Donnell, C.P., 2010. Use of hyperspectral imaging for evaluation of the shelf-life of fresh white button mushrooms (Agaricus bisporus) stored in different packaging films. Innov Food Sci Emerg Technol 11 (3), 423-431. https://doi.org/10.1016/j.ifset.2010.01.016

Tatsika, S., Karamanoli, K., Karayanni, H., Genitsaris, S., 2019. Metagenomic Characterization of Bacterial Communities on Ready-to-Eat Vegetables and Effects of Household Washing on their Diversity and Composition. Pathogens 8, 37. https://doi.org/10.3390/pathogens8010037

Tenenhaus-Aziza, F., Ellouze, M., 2015. Software for predictive microbiology and risk assessment: A description and comparison of tools presented at the ICPMF8 Software Fair. Food Microbiol. 45, 290–299. https://doi.org/10.1016/j.fm.2014.06.026

Torri, L., Sinelli, N., Limbo, S., 2010. Shelf-life evaluation of fresh-cut pineapple by using an electronic nose. Postharvest Biol. Technol. 56, 239–245. https://doi.org/10.1016/j.postharvbio.2010.01.012

Torrione, P., Collins, L.M., Morton, K.D., 2014. Multivariate analysis, chemometrics, and machine learning in laser spectroscopy, in: Baudelet, M. (Ed.), Laser Spectroscopy for Sensing: Fundamentals, Techniques and Applications. Woodhead Publishing Limited, pp. 125–64. https://doi.org/10.1533/9780857098733.1.125

Tournas, V.H., 2005. Spoilage of Vegetable Crops by Bacteria and Fungi and Related Health Hazards. Crit. Rev. Microbiol. 31, 33–44. https://doi.org/10.1080/10408410590886024

Tournas, V.H., Heeres, J., Burgess, L., 2006. Moulds and yeasts in fruit salads and fruit juices. Food Microbiol. 23, 684–688. https://doi.org/10.1016/j.fm.2006.01.003

Trinderup, C.H., Dahl, A., Jensen, K., Carstensen, J.M., Conradsen, K., 2015. Comparison of a multispectral vision system and a colorimeter for the assessment of meat color. Meat Sci. 102, 1–7. https://doi.org/10.1016/j.meatsci.2014.11.012

Truong, V.K., Dupont, M., Elbourne, A., Gangadoo, S., Pathirannahalage, P.R., Cheeseman, S., Chapman, J., Cozzolino, D., 2019. From academia to reality check: A theoretical framework on the use of chemometric in food sciences. Foods 8. https://doi.org/10.3390/foods8050164

Tsakanikas, P., Fengou, L.C., Manthou, E., Lianou, A., Panagou, E.Z., Nychas, G.J.E., 2018. A unified spectra analysis workflow for the assessment of microbial contamination of ready-to-eat green salads: Comparative study and application of non-invasive sensors. Comput. Electron. Agric. 155, 212–219. https://doi.org/10.1016/j.compag.2018.10.025

Tsakanikas, P., Karnavas, A., Panagou, E.Z., 2020. A machine learning workflow for raw food spectroscopic classification in a future industry. Sci Rep 10**,** 11212. https://doi.org/10.1038/s41598-020-68156-2

Tsironi, T., Dermesonlouoglou, E., Giannoglou, M., Gogou, E., Katsaros, G., Taoukis, P., 2017. Shelf-life prediction models for ready-to-eat fresh cut salads : Testing in real cold chain. Int. J. Food Microbiol. 240, 131–140. http://dx.doi.org/10.1016/j.ijfoodmicro.2016.09.032

Unay, D., Gosselin, B., Kleynen, O., Leemans, V., Destain, M.F., Debeir, O., 2011. Automatic grading of Bi-colored apples by multispectral machine vision. Comput. Electron. Agric. 75, 204–212. https://doi.org/10.1016/j.compag.2010.11.006

Ursell, L.K., Metcalf, J.L., Parfrey, L.W. and Knight, R., 2012. Defining the human microbiome. Nutr. Rev. 70 S38-S44. https://doi.org/10.1111/j.1753-4887.2012.00493.x

Valentin-bon, I., Jacobson, A., Monday, S.R., Feng, P.C.H., 2008. Microbiological Quality of Bagged Cut Spinach and Lettuce Mixes. Appl. Environ. Microbiol. 74, 1240–1242. https://doi.org/10.1128/AEM.02258-07

Vassakis, K., Petrakis, E., Kopanakis, I., 2018. Big Data Analytics: Applications, Prospects and Challenges Mobile In: Skourletopoulos, G., Mastorakis, G., Mavromoustakis, C., Dobre, C., Pallis, E. (Eds.), Mobile Big Data, A Roadmap from Models to Technologies 10, Springer International Publishing. https://doi.org/10.1007/978-3-319-67925-9

Ventura-Aguilar, R.I., Colinas-León, M.T., & Bautista-Baños, S., 2017. Combination of sodium erythorbate and citric acid with MAP, extended storage life of sliced oyster mushrooms. LWT 79, 437-444.

Venturini, M.E., Reyes, J.E., Rivera, C.S., Oria, R., & Blanco, D., 2011. Microbiological quality and safety of fresh cultivated and wild mushrooms commercialized in Spain. Food Microbiol. 28, 1492-1498.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naıve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl. Environ. Microbiol. 73, 5261–5267. https://doi.org/10.1128/AEM.00062-07

Wang, H., Peng, J., Xie, C., Bao, Y., He, Y., 2015. Fruit quality evaluation using spectroscopy technology: A review. Sensors 15, 11889–11927. https://doi.org/10.3390/s150511889

Wang, J., Kim, K.H., Kim, S., Kim, Y.S., Li, Q.X., Jun, S., 2010. Simple quantitative analysis of Escherichia coli K-12 internalized in baby spinach using Fourier Transform Infrared spectroscopy. Int. J. Food Microbiol. 144, 147–151. https://doi.org/10.1016/j.ijfoodmicro.2010.09.013

Wang, J., Chen, J., Hu, Y., Hu, H., Liu, G., & Yan, R., 2017. Application of a predictive growth model of Pseudomonas spp. for estimating shelf-life of fresh Agaricus bisporus. J. Food Prot., 80, 1676-1681.

Wang, K., Pu, H., Sun, D.W, 2018. Emerging Spectroscopic and Spectral Imaging Techniques for the Rapid Detection of Microorganisms : An Overview. Compr. Rev. Food Sci. Food Saf. 17. https://doi.org/10.1111/1541-4337.12323

Westad, F., Martens, H., 2000. Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression. J. Near Infrared Spectrosc. 8, 117–124. https://doi.org/10.1255/jnirs.271

Wills, R.B., Golding, J.B., 2016. POSTHARVEST: An introduction to the physiology and handling of fruit and vegetables, sixth ed., UNSW Press, Sydney.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: A basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58, 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Wright, M.N., Ziegler, A., 2017. Ranger: A fast implementation of random forests for high dimensional data in C++ and R. J. Stat. Softw. 77, 1–17. https://doi.org/10.18637/jss.v077.i01

Zagdoun, M., Coeuret, G., N' Dione, M., Champomier-Vergès, M. C., Chaillou, S., 2020. LARGE microbiota survey reveals HOW the microbial ecology of cooked ham is shaped by different processing steps. Food Microbiol 103547. https://doi.org/10.1016/j.fm.2020.103547

Zhang, B.Y., Samapundo, S., Rademaker, M., Noseda, B., Denon, Q., de Baenst, I., Sürengil, G., De Baets, B., Devlieghere, F., 2014. Effect of initial headspace oxygen level on growth and volatile metabolite production by the specific spoilage microorganisms of fresh-cut pineapple. LWT - Food Sci. Technol. 55, 224–231. https://doi.org/10.1016/j.lwt.2013.08.018

Zhang, B., Dai, D., Huang, J., Zhou, J., Gui, Q., Dai, F., 2018. Influence of physical and biological variability and solution methods in fruit and vegetable quality nondestructive inspection by using imaging and near-infrared spectroscopy techniques : A review. Crit Rev Food Sci Nutr 58, No. 12, 2099–2118. https://doi.org/10.1080/10408398.2017.1300789

Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S.G., Alvarez-Cohen, L., 2015. High-Throughput Metagenomic Technologies for Complex Microbial Community Analysis : Open and Closed Formats. mBio 6(1):e02288-146. https://doi.org/10.1128/mBio.02288-14

Zhou, L., Zhang, C., Liu, F., Qiu, Z., He, Y., 2019. Application of Deep Learning in Food: A Review. Compr. Rev. Food Sci. Food Saf. 18, 1793–1811. https://doi.org/10.1111/1541-4337.12492

Zhu, X., Li, S., Shan, Y., Zhang, Z., Li, G., Su, D., Liu, F., 2010. Detection of adulterants such as sweeteners materials in honey using near-infrared spectroscopy and chemometrics. J. Food Eng. 101, 92–97. https://doi.org/10.1016/j.jfoodeng.2010.06.014

Appendix I

Supplementary material (Chapter 3)

Appendix I

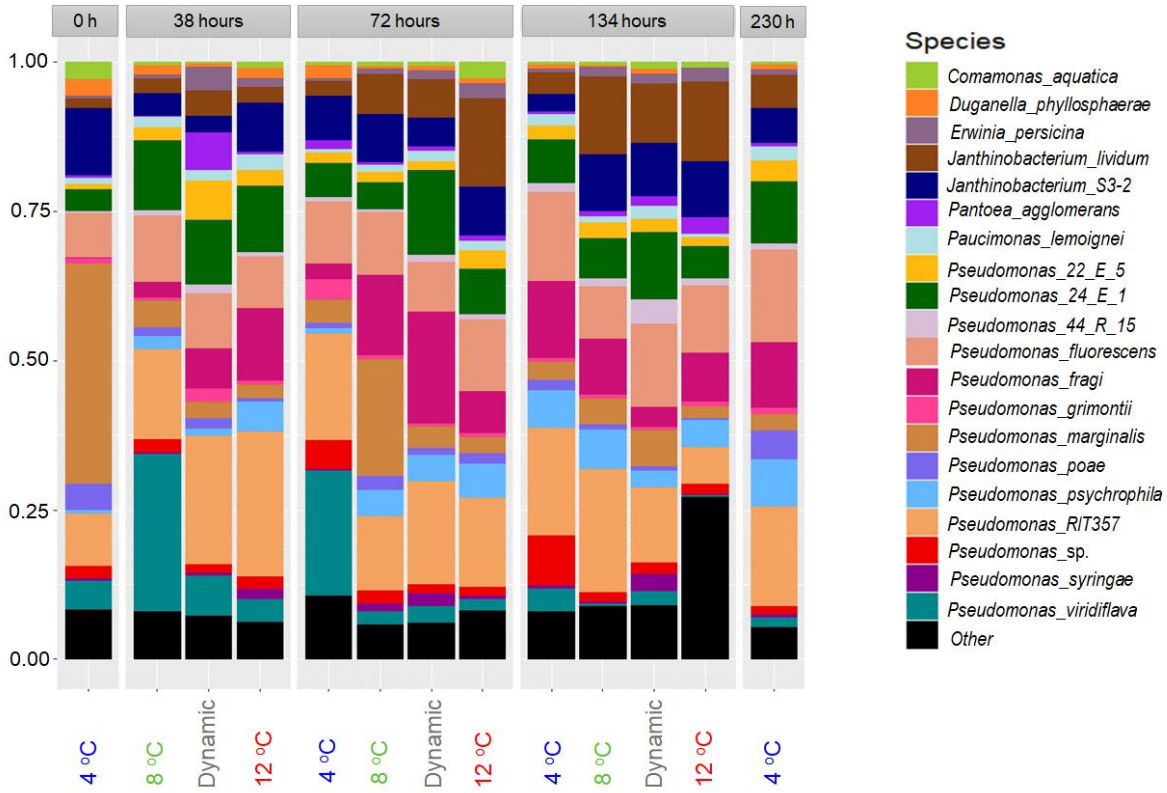


Figure S.3.1. Impact of storage time and temperature on bacterial species composition of ready-to-eat baby spinach samples from Batch BS_1. Samples are clustered according to storage time. Samples are coloured according to storage temperature in blue (4 °C), green (8 °C), red (12 °C) and Dynamic (grey).

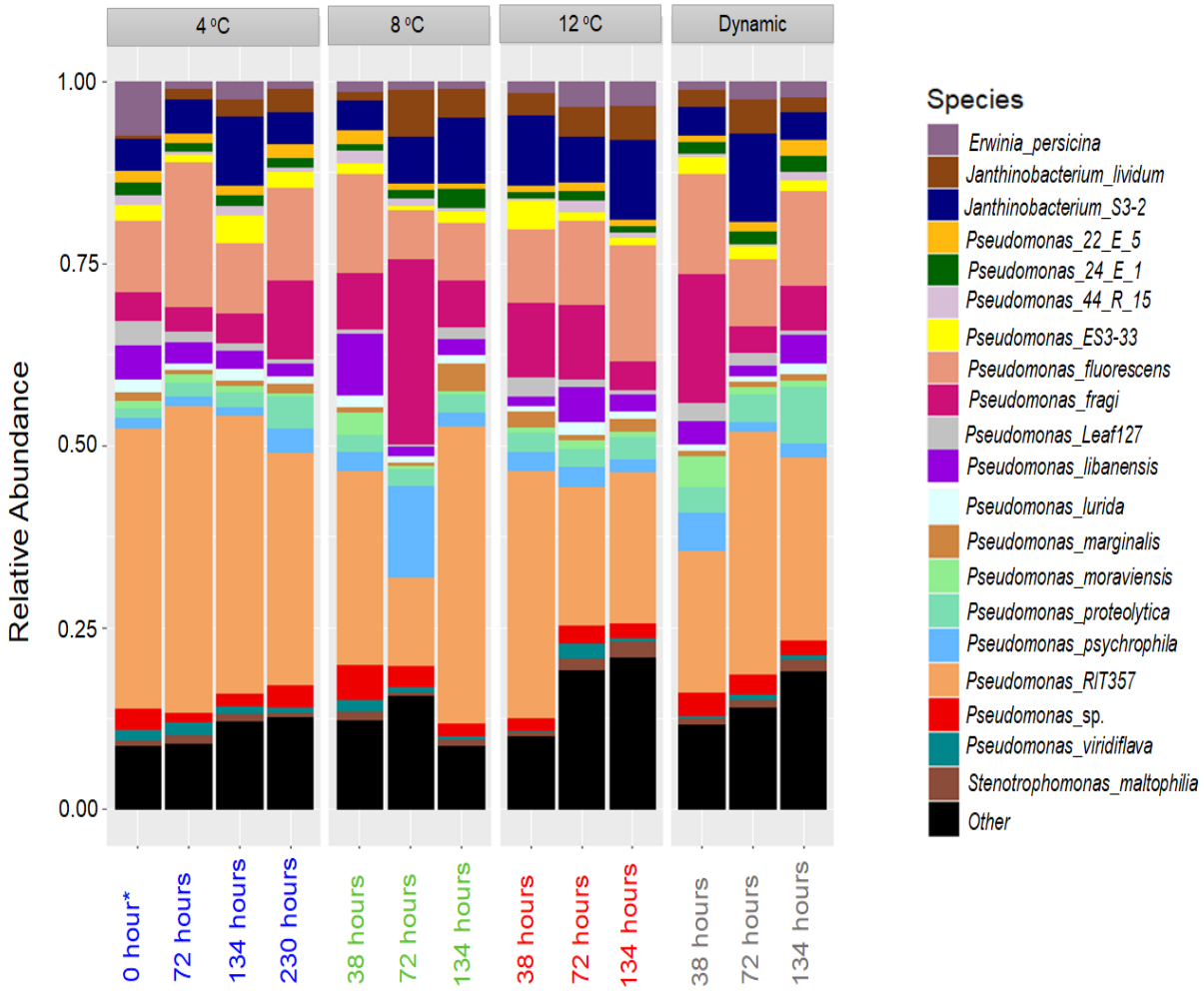


Figure S.3.2. Impact of storage time and temperature on bacterial species composition of ready-to-eat baby spinach samples from Batch BS_2. Samples are clustered according to temperature. Samples are coloured according to storage temperature in blue (4 °C), green (8 °C), red (12 °C) and Dynamic (grey).

Appendix II

Publications and Submissions

Appendix II

Peer reviewed articles

- Tsakanikas, P., Fengou, L.C., **Manthou, E.**, Lianou, A., Panagou, E.Z., Nychas, G.J.E., 2018. A unified spectra analysis workflow for the assessment of microbial contamination of ready-to-eat green salads: Comparative study and application of non-invasive sensors. Comput. Electron. Agric. 155, 212–219. https://doi.org/10.1016/j.compag.2018.10.025

Computers and Electronics in Agriculture 155 (2018) 212–219

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

A unified spectra analysis workflow for the assessment of microbial contamination of ready-to-eat green salads: Comparative study and application of non-invasive sensors

Panagiotis Tsakanikas*, Lemonia-Christina Fengou, Evanthia Manthou, Alexandra Lianou, Efstathios Z. Panagou, George-John E. Nychas*

Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Human Nutrition, School of Food, Biotechnology and Development, Agricultural University of Athens, 11855 Athens, Greece

ARTICLE INFO

Keywords:
Microbiological contamination assessment
Prediction model
Regression
Green salads
Non-invasive sensors

ABSTRACT

The present study provides a comparative assessment of non-invasive sensors as means of estimating the microbial contamination and "time-on-shelf" (i.e. storage time) of leafy green vegetables, using a novel unified spectra analysis workflow. Two types of fresh ready-to-eat green salads were used in the context of this study for the purpose of evaluating the efficiency and practical application of the presented workflow: rocket and baby spinach salads. The employed analysis workflow consisted of robust data normalization, powerful feature selection based on random forests regression, and selection of the number of partial least squares regression coefficients in the training process by estimating the "knee-point" on the explained variance plot. Training processes were based on microbiological and spectral data derived during storage of green salad samples at isothermal conditions (4, 8 and 12 °C), whereas testing was performed on data during storage under dynamic temperature conditions (simulating real-life temperature fluctuations in the food supply chain). Since an increasing interest in the use of non-invasive sensors in food quality assessment has been made evident in recent years, the unified spectra analysis workflow described herein, by being based on the creation/usage of limited sized featured sets, could be very useful in food-specific low-cost sensor development.

1. Introduction

In recent years, the production, sale and consumption of fresh raw fruits and vegetables, especially as pre-cut and ready-to-eat (minimally processed by washing, slicing or shredding and packaging) salads have undergone substantial increases in the European Union as well as the United States. This is the result of consumer preferences for fresher, more convenient and nutritious foods that meet the needs of busier lifestyles, at least in the developed countries. Vegetables are highly contaminated by nature due either to agronomic systems employed in their production (e.g., irrigation with contaminated water, organic fertilizers such as manure, etc.) or during processing, handling, and marketing, (Francis et al., 2012; Skandamis and Nychas, 2011). Nonetheless, consumers demand fresh produce commodities, which should be not only perfectly safe for human consumption, but visually attractive as well.

Over the past decade, a number of sensors based on vibrational spectroscopy or hyperspectral/multispectral imaging have been developed (Rateni et al., 2017). Although such sensors have started gaining popularity as rapid and efficient methods for assessing food quality and/or food composition, their utilization in handling of microbiological issues related to the current EU legislation (Commission, 2005) still remains to be firmly established. Indeed, the challenge of using non-invasive sensors as sensible alternatives to the costly and time-consuming conventional microbiological techniques has not been adequately tackled yet.

Spectroscopy which indeed allows rapid, non-invasive and non-destructive analysis (Nychas et al., 2016), cannot be considered as target-specific method and it must be associated with complex data analytics to extract the relevant microbiological and biochemical information. This can be achieved by the selections of the wavelength range for source and detector, and of the measurement setup. It is evident that due to the multi-dimensional nature of the data generated from such analyses, the output needs to be coupled with a suitable statistical

* Corresponding authors.
E-mail addresses: p.tsakanikas@aua.gr (P. Tsakanikas), lefengou@gmail.com (L.-C. Fengou), evita.m@windowslive.com (E. Manthou), alianou@aua.gr (A. Lianou), stathispanagou@aua.gr (E.Z. Panagou), gjn@aua.gr (G.-J.E. Nychas).

https://doi.org/10.1016/j.compag.2018.10.025
Received 4 August 2018; Received in revised form 15 October 2018; Accepted 16 October 2018

- **Manthou, E.**, Tarlak, F., Lianou, A., Ozdemir, M., Zervakis, G. I., Panagou, E. Z., & Nychas, G. J. E., 2019. Prediction of indigenous Pseudomonas spp. growth on oyster mushrooms (*Pleurotus ostreatus*) as a function of storage temperature. *LWT*, 111, 506-512. https://doi.org/10.1016/j.lwt.2019.05.062

LWT - Food Science and Technology 111 (2019) 506–512

Contents lists available at ScienceDirect

LWT - Food Science and Technology

journal homepage: www.elsevier.com/locate/lwt

Prediction of indigenous *Pseudomonas* spp. growth on oyster mushrooms (*Pleurotus ostreatus*) as a function of storage temperature

Evanthia Manthou[a,1], Fatih Tarlak[b,1], Alexandra Lianou[a], Murat Ozdemir[c], Georgios I. Zervakis[d], Efstathios Z. Panagou[a], George-John E. Nychas[a,*]

[a] Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Human Nutrition, School of Food, Biotechnology and Development, Agricultural University of Athens, 11855, Athens, Greece
[b] Department of Nutrition and Dietetics, Istanbul Gedik University, 34876, Kartal, Istanbul, Turkey
[c] Department of Chemical Engineering, Gebze Technical University, 41400, Gebze, Kocaeli, Turkey
[d] Laboratory of General and Agricultural Microbiology, Department of Crop Science, School of Agricultural Production, Infrastructure and Environment, Agricultural University of Athens, 11855, Athens, Greece

ARTICLE INFO

Keywords:
Growth kinetics
Microbial spoilage
Modelling
Storage experiments

ABSTRACT

The growth kinetic behaviour of *Pseudomonas* spp. naturally occurring on oyster mushrooms (*Pleurotus ostreatus*) was evaluated during storage at different isothermal conditions (4, 10 and 16 °C), and was described quantitatively using a one-step global parameter estimation method. In the context of this modelling approach, the growth kinetic parameters of maximum specific growth rate (μ_{max}) and lag phase duration (λ) were estimated using the Baranyi model, whereas the effect of temperature on μ_{max} was described using a secondary square-root-type model. The global model's goodness-of-fit indices of root mean square error (RMSE) and adjusted coefficient of determination (adjusted-R^2) were estimated to be 0.206 and 0.948, respectively. The global model was then externally validated using growth data generated during storage of oyster mushrooms under dynamic temperature conditions. Specifically, the differential form of the Baranyi model merged with the square-root-type model was solved numerically using the fourth-order Runge-Kutta method in order to predict the *Pseudomonas* spp. concentration on mushrooms under fluctuating temperature conditions. The developed dynamic modelling approach exhibited satisfactory performance, with the mean deviation and the mean absolute deviation being −0.10 and 0.22 log CFU/g, respectively. Along with further substantiation and optimization, the developed model should be useful in food quality management systems, aiming in particular at the improvement of the microbiological quality of oyster mushrooms.

1. Introduction

Edible mushrooms are known as nutrient-rich food commodities containing high amounts of proteins, minerals and bioactive compounds (Wani, Bodha, & Wani, 2010). At the same time, due to their high moisture content and neutral pH, fresh mushrooms constitute an ideal substrate for microbial growth (Venturini, Reyes, Rivera, Oria, & Blanco, 2011). Although numerous edible mushroom species exist in nature, oyster mushroom (*Pleurotus ostreatus*) is one of the most commonly produced and consumed species in the world (Venturini et al., 2011). Bacterial species belonging to the genus *Pseudomonas* tend to dominate on oyster mushrooms (Reyes, Venturini, Oria, & Blanco, 2004; Venturini et al., 2011), with the microbial spoilage of these food commodities proceeding fast and their commercial value being lost within a few days from production (Ventura-Aguilar, Colinas-León, & Bautista-Baños, 2017).

Conventional microbiological and/or physicochemical methods applied in the framework of food storage experiments, frequently under abusive environmental conditions, have been extensively used by both the food industry and academia for the purpose of shelf-life assessment. Nonetheless, since several limitations of this approach have been identified, its utilization for the accurate and reproducible estimation of a food product's shelf-life was early treated with scepticism by the scientific community (McMeekin & Ross, 1996). For instance, the estimation of shelf-life based on this approach is valid only for the conditions tested, while any changes to these conditions require repetition of the storage experiments. Furthermore, no information is provided on the magnitude of influence of the controlling factors on microbial

* Corresponding author.
E-mail address: gjn@aua.gr (G.-J.E. Nychas).
[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.lwt.2019.05.062
Received 26 December 2018; Received in revised form 5 April 2019; Accepted 12 May 2019
Available online 14 May 2019
0023-6438/

- **Manthou, E.**, Lago, S.L., Dagres, E., Lianou, A., Tsakanikas, P., Panagou, E.Z., Anastasiadi, M., Mohareb, F., Nychas, G.J.E., 2020. Application of spectroscopic and multispectral imaging technologies on the assessment of ready-to-eat pineapple quality : A performance evaluation study of machine learning models generated from two commercial data analytics tools. Comput. Electron. Agric. 175, 105529. https://doi.org/10.1016/j.compag.2020.105529

Computers and Electronics in Agriculture 175 (2020) 105529

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Application of spectroscopic and multispectral imaging technologies on the assessment of ready-to-eat pineapple quality: A performance evaluation study of machine learning models generated from two commercial data analytics tools

Evanthia Manthou[a], Sergio-Llaneza Lago[b], Evaggelos Dagres[a], Alexandra Lianou[a], Panagiotis Tsakanikas[a], Efstathios Z. Panagou[a], Maria Anastasiadi[b], Fady Mohareb[b], George–John E. Nychas[a,*]

[a] Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Human Nutrition, School of Food & Nutritional Sciences, Agricultural University of Athens, 11855 Athens, Greece
[b] Bioinformatics Group, Department of Agrifood, School of Water, Energy and Environment Cranfield University, College Road, Cranfield, Bedfordshire MK43 0AL, UK

ARTICLE INFO

Keywords:
Pineapple
Quality
Vibrational spectroscopy
Multispectral imaging
Machine learning
SorfML

ABSTRACT

Recently, rapid, non-invasive analytical methods relying on vibrational spectroscopy and hyper/multispectral imaging, are increasingly gaining popularity in food science. Although such instruments offer a promising alternative to the conventional methods, the analysis of generated data demands complex multidisciplinary approaches based on data analytics tools utilization. Therefore, the objective of this work was to (i) assess the predictive power of different analytical platforms (sensors) coupled with machine learning algorithms in evaluating quality of ready-to-eat (RTE) pineapple (*Ananas comosus*) and (ii) explore the potentials of The Unscrambler software and the online machine-learning ranking platform, SorfML, in developing the predictive models required by such instruments to assess quality indices. Pineapple samples were stored at 4, 8, 12 °C and dynamic temperatures and were subjected to microbiological (total mesophilic microbial populations, TVC) and sensory analysis (colour, odour, texture) with parallel acquisition of spectral data. Fourier-transform infrared, fluorescence (FLUO) and visible sensors, as well as Videometer instrument were used. For TVC, almost all the combinations of sensors and Partial-least squares regression (PLSR) algorithm from both analytics tools reached values of root mean square error of prediction (RMSE) up to 0.63 log CFU/g, as well as the highest coefficient of determination values (R^2). Moreover, Linear Support Vector Machine (SVM Linear) combined with each one of the sensors reached similar performance. For odour, FLUO sensor achieved the highest overall performance, when combined with Partial-least squares discriminant analysis (PLSDA) in both platforms with accuracy close to 85%, but also with values of sensitivity and specificity above 85%. The SVM Linear and MSI combination also achieved similar performance. On the other hand, all models developed for colour and texture showed poor prediction performance. Overall, the use of both analytics tools, resulted in similar trends concerning the feasibility of the different analytical platforms and algorithms on quality evaluation of RTE pineapple.

1. Introduction

In the context of tremendous technological change, a growing lack of natural resources, and a continuous evolution of consumers' lifestyles and consumption habits across the globe, food industry is challenged to provide safe and qualitative food to consumers. To address the need for efficient, safe and environmental respectful production, as well as strict communication and connection with the consumers, several approaches have been developed (Nychas et al., 2016).

Among these, analytical methods based on vibrational spectroscopy and hyperspectral/multispectral imaging have gained the attention of scientists, since they could fulfill the needs of food industry as rapid and efficient methods for assessing food quality (Fengou et al., 2019b; Barbin et al., 2015; Papadopoulou et al., 2011; Ammor et al., 2009; Camps and Christen, 2009), safety (Grewal et al., 2015; Brandily et al., 2011; Davis et al., 2010; Wang et al., 2010) and authentication-

* Corresponding author.
E-mail address: gjn@aua.gr (G.E. Nychas).

https://doi.org/10.1016/j.compag.2020.105529
Received 2 May 2020; Received in revised form 24 May 2020; Accepted 24 May 2020

- Argyri, K., Doulgeraki, A. I., **Manthou, E.**, Grounta, A., Argyri, A. A., Nychas, G-J E., Tassou, C.C., 2020. Microbial Diversity of Fermented Greek Table Olives of Halkidiki and Konservolia Varieties from Different Regions as Revealed by Metagenomic Analysis. Microorganisms 8, 1241. doi:10.3390/microorganisms8081241

microorganisms

MDPI

Article

Microbial Diversity of Fermented Greek Table Olives of Halkidiki and Konservolia Varieties from Different Regions as Revealed by Metagenomic Analysis

Konstantina Argyri [1], Agapi I. Doulgeraki [1,*], Evanthia Manthou [2], Athena Grounta [1], Anthoula A. Argyri [1], George-John E. Nychas [2] and Chrysoula C. Tassou [1,*]

[1] Institute of Technology of Agricultural Products, Hellenic Agricultural Organisation DEMETER, Sofokli Venizelou 1, Lycovrissi, 14123 Athens, Greece; nargiri@gmail.com (K.A.); athenagrounta@gmail.com (A.G.); anthi.argyri@gmail.com (A.A.A.)

[2] Laboratory of Food Microbiology and Biotechnology, Department of Food Science and Human Nutrition, School of Food and Nutritional Sciences, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece; evita.m@windowslive.com (E.M.); gjn@aua.gr (G.-J.E.N.)

* Correspondence: adoulgeraki@aua.gr (A.I.D.); ctassou@nagref.gr (C.C.T.); Tel.: +30-2102845940 (A.I.D. & C.C.T.)

Received: 4 August 2020; Accepted: 13 August 2020; Published: 14 August 2020

check for updates

Abstract: Current information from conventional microbiological methods on the microbial diversity of table olives is insufficient. Next-generation sequencing (NGS) technologies allow comprehensive analysis of their microbial community, providing microbial identity of table olive varieties and their designation of origin. The purpose of this study was to evaluate the bacterial and yeast diversity of fermented olives of two main Greek varieties collected from different regions—green olives, cv. Halkidiki, from Kavala and Halkidiki and black olives, cv. Konservolia, from Magnesia and Fthiotida—via conventional microbiological methods and NGS. Total viable counts (TVC), lactic acid bacteria (LAB), yeast and molds, and *Enterobacteriaceae* were enumerated. Microbial genomic DNA was directly extracted from the olives' surface and subjected to NGS for the identification of bacteria and yeast communities. *Lactobacillaceae* was the most abundant family in all samples. In relation to yeast diversity, *Phaffomycetaceae* was the most abundant yeast family in Konservolia olives from the Magnesia region, while *Pichiaceae* dominated the yeast microbiota in Konservolia olives from Fthiotida and in Halkidiki olives from both regions. Further analysis of the data employing multivariate analysis allowed for the first time the discrimination of cv. Konservolia and cv. Halkidiki table olives according to their geographical origin.

Keywords: table olives; Halkidiki olives; Konservolia olives; NGS; Greek-style fermentation; Spanish-style fermentation; microbiological analysis; metagenomic analysis

1. Introduction

Table olives are an important fermented food in Mediterranean countries with great nutritional and economic significance. Their content in bioactive compounds, vitamins, dietary fibers, unsaturated fatty acids, minerals, and antioxidants with demonstrated positive effects on human health meets the consumers' needs toward natural or minimal processed foods that, beyond basic nutrition, offer additional health benefits [1]. Raised awareness of the health benefits of olives may be partially the driving force for the increased global table olive consumption that has doubled over the past three decades and is expected to increase by 2.1 percent in 2020, as predicted by the International Olive Council (IOC) [2].

- **Manthou, E.**, Coeuret, G., Chaillou, S., Nychas, G. J. E., 2021. Evolution of fungal community associated with ready-to-eat pineapple during storage under different temperature conditions. Food Microbiol. 97, 103736. https://doi.org/10.1016/j.fm.2021.103736

Food Microbiology 97 (2021) 103736

Contents lists available at ScienceDirect

Food Microbiology

journal homepage: www.elsevier.com/locate/fm

Evolution of fungal community associated with ready-to-eat pineapple during storage under different temperature conditions

Evanthia Manthou [a], Gwendoline Coeuret [b], Stephane Chaillou [b, *], George-John E. Nychas [a, **]

[a] Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Human Nutrition, Agricultural University of Athens, Iera Odos 75, 11855, Athens, Greece
[b] Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, F78352, Jouy-en-Josas, France

ARTICLE INFO

Keywords:
Ready-to-eat pineapple
Spoilage
Fungal diversity
Metagenetic analysis
Temperature and time effect

ABSTRACT

The international market of fresh-cut products has witnessed dramatic growth in recent years, stimulated by consumer's demand for healthy, nutritious and convenient foods. One of the main challenging issues for the quality and safety of these products is the potential microbial spoilage that can significantly reduce their shelf-life. The complete identification of fresh-cut product microbiota together with the evaluation of environmental factors impact on microbial composition is of primary importance. We therefore assessed the fungal communities associated with the spoilage of ready-to-eat (RTE) pineapple using a metagenetic amplicon sequencing approach, based on the ITS2 region. Our results revealed a significant variability on fungal species composition between the different batches of RTE pineapple. The initial microbiota composition was the main influencing factor and determined the progress of spoilage. Temperature and storage time were the secondary factors influencing spoilage and their impact was depending on the initial prevalent fungal species, which showed different responses to the various modifications. Our results strongly suggest that further large-scale sampling of RTE pineapple production should be conducted in order to assess the full biodiversity range of fungal community involved in the spoilage process and for unravelling the impact of important environmental factors shaping the initial microbiota.

1. Introduction

Fresh-cut market has grown dramatically in recent years, as a result of changes on consumers' attitude. Ready-to-eat (RTE) fruits and vegetables fulfil the growing demand for healthy, convenient and minimally processed food products (Gorni et al., 2015; Qadri et al., 2015). However, the quality and safety assurance of these new types of fresh products is a major challenge for the fresh-cut industry and requires full involvement and increasing investigations of food scientists (Padrón-Mederos et al., 2020).

Fresh-cut fruits and vegetables products have a limited shelf life due to accelerated physiological and biochemical changes occurring during their processing and storage (Di Egidio et al., 2009; Torri et al., 2010; Zhang et al., 2014). Indeed, the processing treatments render the products more prone to spoilage microorganisms, as well as micro-organisms of public health significance (Leff and Fierer, 2013; Qadri et al., 2015). Various studies underline the presence of phytopathogens and human pathogens, but also microorganisms with antagonistic properties against these pathogens, which have a significant influence on human health and products' quality (Gorni et al., 2015). Therefore, a better insight into the microbial community and its potential interactions in food associated matrices is required to provide safe and high-quality food (Juste et al., 2008; Cao et al., 2017).

So far culture-dependent methods have been the gold standards in food microbiology, since they have led to the description of a number of habitats. However, they are extremely biased in their ability to unravel the microbial communities of complex matrices associated with food or environmental samples (Juste et al., 2008; Ercolini, 2013; Zhou et al., 2015; Edet et al., 2017). On the other hand, the development of next generation sequencing (NGS) techniques has enabled researchers to study food microbial ecology from broader and deeper perspectives. Recently, metagenetic and metagenomic approaches have resulted in improved understanding of a microbiota by providing a species- and strain-level characterization (Abdelfattah et al., 2016, 2018; Poirier

* Corresponding author.
** Corresponding author.
E-mail addresses: gjn@aua.gr (S. Chaillou), stephane.chaillou@inrae.fr (G.-J.E. Nychas).

https://doi.org/10.1016/j.fm.2021.103736
Received 26 September 2020; Received in revised form 5 January 2021; Accepted 6 January 2021
Available online 19 January 2021

- **Manthou**, **E.**, Karnavas, A., Fengou, L.-C., Bakalaki, A., Lianou, A., Tsakanikas, P., & Nychas, G. J. E. Spectroscopy and imaging technologies coupled with machine learning for the assessment of the microbiological spoilage associated to ready-to-eat leafy vegetables. ***Submitted in International Journal of Food Microbiology***

International conferences

- **Manthou, E.**, Dagres, E., Lianou, A., Nychas, G.J.E, 2018. Microbiological Spoilage of Cut Ready-to-Eat Pineapple during Storage under Different Temperature Conditions. 2018 European Symposium. IAFP's European Symposium on Food Safety, 25 – 27 April 2018 in Stockholm, Sweden. (poster presentation)
- Tsakanikas, P., **Manthou, E.**, Fengou, L.-C., Lianou, A., Panagou, E.Z., Nychas, G.J.E, 2018. Spectroscopy-based sensors under a unified feature selection approach for microbial contamination and storage time prediction of ready-to-eat rocket. Food Micro 2018 (FM 2018): 3-6 September 2018, Germany. (poster presentation)
- Tarlak, F., **Manthou, E.**, Dagres, E., Lianou, A., Murat Ozdemir, Nychas, G.J.E, 2018. Modelling the growth kinetics of *Pseudomonas* spp. on *Pleurotus ostreatus* mushrooms under non-isothermal conditions. Food Micro 2018 (FM 2018): 3-6 September 2018, Germany. (poster presentation)
- **Manthou, E.**, Dagres, E., Lianou, A., Tsakanikas, P., Panagou, E.Z., Nychas, G.J.E, 2019. Application of Fluorescence Spectroscopy as a Tool for Microbial Spoilage Assessment in Fresh-cut Pineapple. IAFP 2019 Annual Meeting, July 21-24 2019 in Louisville, Kentucky. (poster presentation)
- **Manthou**, E., Bakalaki, A., Lianou, A., Tsakanikas, P., Panagou, E.Z., Nychas, G.J.E, 2019. Estimation of Microbial Spoilage of Ready-to-eat Baby Spinach using Fourier Transform Infrared Spectroscopy. 2019 European Symposium. IAFP's European Symposium on Food Safety 24 -26 April 2019 in Nantes, France. (oral presentation)
- **Manthou**, **E.,** Coeuret, G., Nychas, G.-J. E., Chaillou, S., 2019. Evolution of microbial community associated with ready-to-eat fruits and vegetables during storage under

different temperature conditions. 5th International Conference on Microbial Diversity 2019, 25-27 September 2019 in Catania, Italy. (oral presentation)

- **Manthou**, E., Kochila, A., Tzamourani, A., Chaillou, S., Nychas, G.J.E., Panagou, E., 2021. A metagenetic analysis of bacterial community in inoculated fermentations of Conservolea variety black olives with multifunctional starter cultures in reduced salt brines. 2021 European Symposium. IAFP's European Symposium on Food Safety 27 -28 April 2021 (accepted abstract), virtual meeting. (poster presentation)

National conferences

- **Manthou E**., Dagres E., Bonatsou S., Doulgeraki A.I., Nychas G.-J.E., 2019. Molecular characterization of *Pseudomonas* spp.isolatesfrom sporocarps of oyster mushrooms with yellow blotch disease. 8th MBK conference, 18-20 April 2019 in Patra, Greece. (poster presentation)
- **Μάνθου**, **Ε.,** Δαγρές, Ε., Λιανού, Α., Τσακανίκας, Π., Πανάγου, Ε.Ζ. , Νυχάς, Γ.-Ι., 2019. Εκτίμηση της μικροβιολογικής αλλοίωσης έτοιμου προς κατανάλωση ανανά μέσω της φασματοσκοπίας υπερύθρου και της πολυφασματικής απεικόνισης. 12ο Πανελλήνιο Επιστημονικό Συνέδριο Χημικής Μηχανικής, 29-31 Μαΐου 2019 στην Αθήνα, Ελλάδα. (poster presentation)