Subject of studies: M.Sc Biochemistry
Bachelor/Master/State exam: Master
Time of internship: 13.10.2025 – 19.12.2025
Place of internship: Paris, France
Internship institution: INSERM Paris

1. Planning and Preparation

1.1 Application process of Internship

I usually searched for internships not by applying through official programs, but by directly emailing principal investigators whose research interests aligned with mine. I remember sending emails to around eight laboratories in France. Among them, I was contacted by INSERM Paris, and after a 40-minute interview, the internship was confirmed.

1.2 Finding Accomodation

Paris is a highly international city. One interesting aspect is that different areas tend to have different ethnic communities concentrated in them. Since I was going to stay for only about two and a half months, I decided to look for a sublet. As a Korean, I focused on finding housing through Korean community networks, which ultimately helped me secure accommodation. I lived in the 13th arrondissement for about three weeks and in the 15th arrondissement for around seven weeks. If you do not speak French or are staying for a short period of time, I would recommend looking for housing through a community related to your own background.

2. Internship

During my internship period (10 weeks), I was mainly involved in three types of tasks.

2.1 Data analysis

2.1.1 Methods Subsection

Raw SRA sequences from NCBI BioProjects were downloaded using SRA Toolkit after generating accession lists of SRR/ERR identifiers. Files were renamed for single-end or paired-end data, quality-trimmed with Fastp, and paired reads merged using FLASH with parameters adjusted for read length (e.g., --max-overlap 250 for 250 bp V4 regions). Fastq files were converted to fasta format using seqkit, and sequences were BLASTed against a local 16S database (SFBblastdb.fna) at 94% identity and 90% query coverage with multi-threading.

Positive SFB hits were extracted by merging BLAST outputs, identifying unique query IDs, and retrieving corresponding sequences into HITS.fasta, with short reads filtered. Read counts were quantified per sample, metadata extracted and merged from BioProject files, and SFB frequencies calculated as (positive reads / total reads) × 100.

2.1.2 Results Subsection

A total of 24 projects were successfully completed. In addition, 8 projects remained uncompleted initially. The main reason for the incomplete projects was troubleshooting difficulties encountered in the early stages of data processing. Specifically, challenges arose during the handling of raw data and the execution of positive and negative transfer processes, which temporarily hindered progress. These issues were identified and systematically resolved in subsequent attempts, ensuring that the workflow could be properly followed for future projects. This experience highlights the importance of early-stage data validation and troubleshooting in complex analyses.

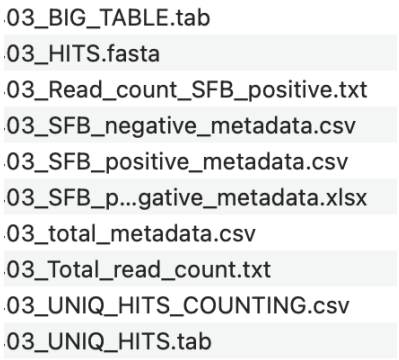

Figure 1. Analysis of Projects; Examples; Using metadata obtained from NCBI and ENA, SFB samples were classified as positive or negative

2.2 Wet-lab experiments (with PhD student S.)

2.2.1 Methods Subsection

Bacterial Transformation Workflow: Collaborative experiments were conducted with PhD student S. to establish a complete bacterial transformation protocol. The workflow encompassed PCR amplification, DNA extraction, and bacterial competency assays, with molecular characterization at each step via Western blotting and DNA analysis. During these experiments, PCR amplification successfully produced target products, specifically GST and pSTE019 fragments.

Western blot: Protein samples were separated by SDS-PAGE and transferred to PVDF membranes. Primary antibody incubation was performed using a single-day protocol at room temperature for rabbit antibodies or a two-day protocol with overnight incubation at 4°C for optimized binding. Following incubation, membranes underwent six sequential washing steps using TBST to remove unbound antibodies, then HRP-conjugated secondary antibodies were applied for chemiluminescent detection. Antibody purity was independently verified using parallel Coomassie Brilliant Blue staining. Subsequent experiments will assess GFP or FLAG-tagged protein expression (<50 kDa) in four mammalian cell lines under induced and non-induced conditions at two time points (16 total samples),

with vinculin as a loading control. Additionally, recombinant proteins expressed in *E. Coli* will be analyzed for expression level and solubility by examining both soluble and insoluble fractions.

PCR Amplification and Product Generation: PCR amplification was performed to generate target constructs, specifically GST and pSTE019 fragments.

DNA Extraction and Quality Control: Plasmid DNA extraction from pSTE019 was performed using DNA extraction Kit (BioLab). Extracted DNA was quantified spectrophotometrically, yielding a concentration of 68.6 ng/uL. The A260/280 ratio of 1.86 indicates high purity and genomic integrity, meeting acceptance criteria for downstream applications (typical threshold: >1.8)

2.2.2 Results Subsection

Following primary antibody incubation, Lane A displayed an exceptionally intense and well-defined band, indicating high target protein concentration and excellent antibody specificity. During the sequential washing phase, lanes W1 and W2 retained substantial signal intensity, reflecting antibodies selectively bound to specific epitopes, while lanes W3–W6 showed progressively diminished signal as non-specifically bound antibodies were removed (Figure 2A).

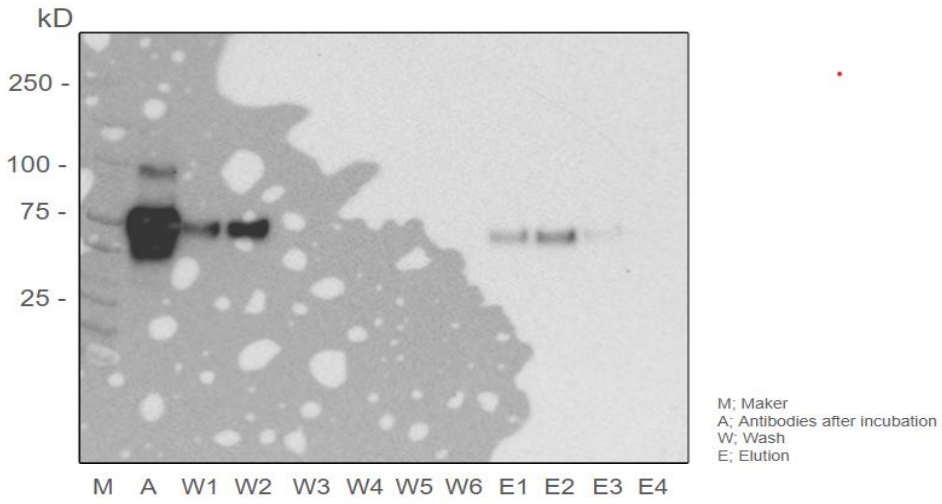

Figure 2A. Western Blot; for detection of Antibodies.

The elution fractions revealed successful protein recovery with high specificity. Lanes E1 and E2 exhibited clear, sharply defined single bands, confirming effective concentration and isolation of the target protein. Lane E3 showed markedly weaker signal reflecting progressive protein depletion, and lane E4 displayed barely detectable signal indicating near-complete depletion. This pattern validates the antibody's high specificity and the purification method's efficiency for subsequent cell-based and bacterial protein expression studies, with parallel Coomassie staining confirming antibody purity.

PCR Amplification Success: Both GST and pSTE019 target fragments were successfully amplified, producing clean, single-band products on agarose gel electrophoresis (Figure 2B).

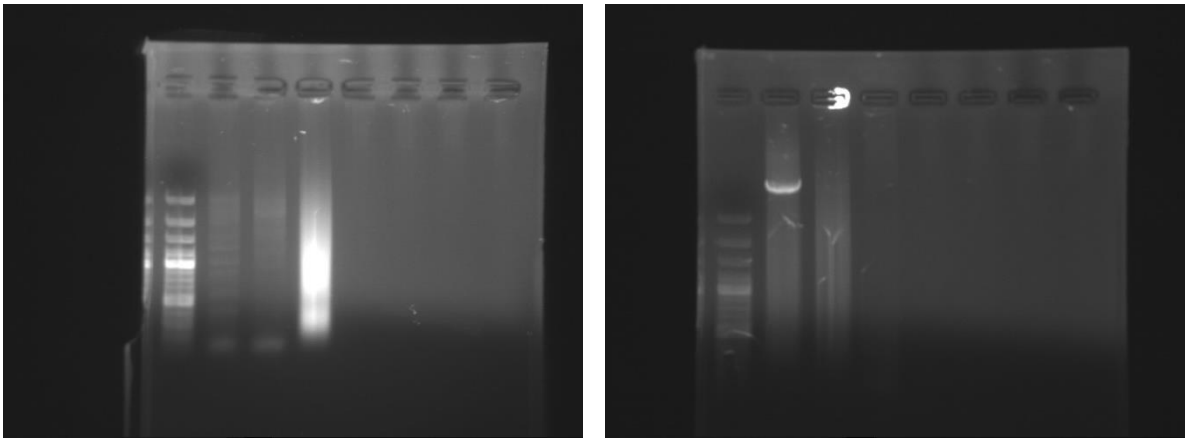

Figure 2B. pSTE109 PCR product ; First trial (left), Second trial (right)

DNA Quality Metrics: The A260/280 absorbance ratio of 1.86 is consistent with protein-free, high-quality DNA. This quality threshold is critical for successful transformation efficiency in downstream cloning procedures.

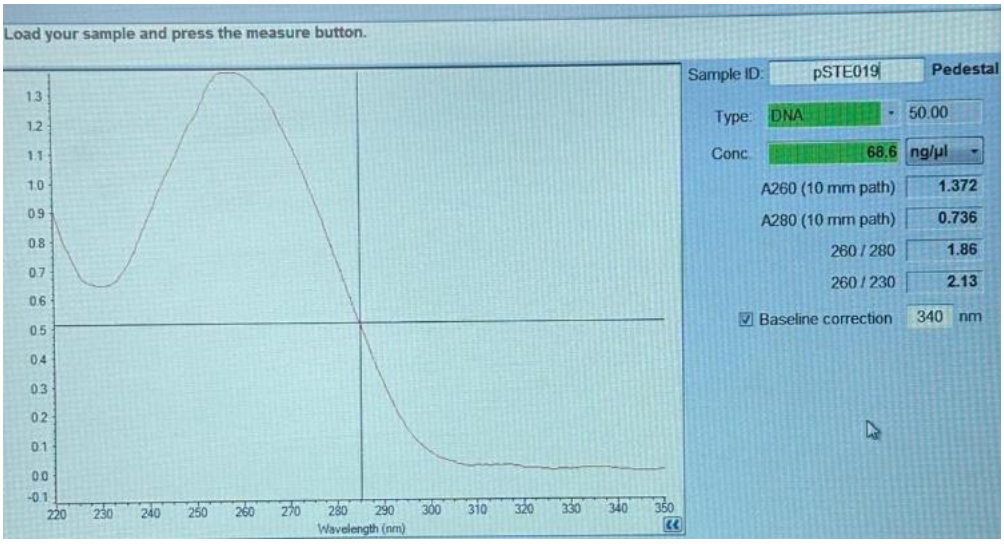

Figure 2C. pSTE019 DNA extraction

Transformation Efficiency Analysis: Despite successful PCR amplification and DNA quality confirmation, bacterial transformation results in no colonies. Since upstream molecular characterization (PCR and DNA extraction) yielded expected results, the transformation failure likely originated from downstream steps, including: (1) competent cell preparation or viability assessment, (2) electroporation/chemical transformation parameters, (3) recovery media conditions, or (4) antibiotic selection stringency.

2.3 Bacteria Morphology analysis

2.3.1 Methods Subsection

Morphological Measurement and Calibration: Segmented filamentous bacteria (SFB) were imaged across 20 biological/technical replicates by Falcon SFB. Image analysis was conducted using Fiji. Measurements were calibrated using embedded scale bars with known increments (5 um and 10 um unit divisions), enabling standardized quantification across image sets. Bacterial diameter was measured at multiple points per bacteria to ensure reproducibility.

Statistical Analysis: Morphological variations were categorized into diameter. Diameter measurements were correlated with morphological phenotypes using simple-linear correlation.

2.3.2 Results Subsection

The diameters of segmented filamentous bacteria (SFB) were analyzed across 20 image files. The results indicated that most changes were statistically significant, demonstrating a clear correlation between morphological changes and bacterial diameter. This analysis provides quantitative evidence supporting the impact of morphological differences on SFB size.

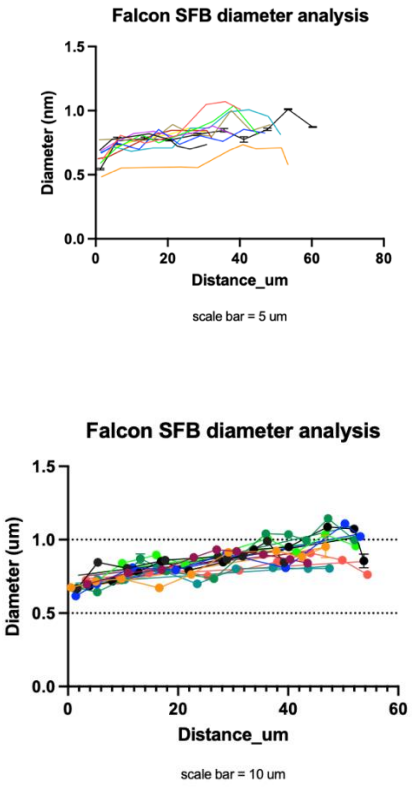

Figure 3. Falcon SFB diameter analysis; 5 um scale bar (top) and 10 um scale bar (bottom)

Correlation analysis: Statistical analysis revealed a significant correlation between bacterial morphology and diameter.

2.4. Discussion

The internship integrated computational, molecular, and microscopic approaches to microbiome research. The successful completion of 24 bioinformatics projects after resolving initial troubleshooting challenges in data processing underscores the importance of systematic validation in high-throughput sequence analysis. Upstream molecular characterization—PCR amplification producing clean products and plasmid DNA achieving high-quality metrics (A260/280 = 1.86)—demonstrated technical competency, though bacterial transformation yielded no colonies, suggesting downstream steps rather than DNA preparation require optimization. The Western blot analysis exhibited excellent specificity with progressive signal diminishment across washing steps, validating antibody-epitope interactions and purification efficiency. Morphological analysis of segmented filamentous bacteria revealed statistically significant correlations between diameter and phenotype, providing quantitative evidence for morphological variation. Collectively, these results illustrate both technical achievements and the diagnostic value of systematic troubleshooting in refining research protocols.

2.5 Conclusion

This ten-week internship provided comprehensive exposure to integrated microbiome research across bioinformatics, molecular biology, and microscopy. The successful SFB classification of 24 projects combined with developed bacterial transformation and Western blot protocols demonstrates feasibility of multi-scale analytical approaches. Though bacterial transformation remains to be optimized, the high-quality upstream characterization and exceptional Western blot specificity indicate sound experimental design. The quantitative morphological analysis establishing significant SFB diameter-phenotype correlations provides a foundation for subsequent investigations. This experience reinforced the importance of iterative troubleshooting and integrated experimental design while acquiring transferable skills in bioinformatics, molecular biology, and image analysis applicable to contemporary biochemical research environments.

2.6 Lab life

In addition to working on these three main tasks, I attended seminars with members of the lab as well as collaborative seminars involving other research institutes.

3 Social Contacts

In this lab, the PI frequently interacted with students, so discussions about experimental results were held flexibly, not only on the scheduled days but also as needed. Regarding networking events, the institute is located near the Pasteur Institute, and collaborative seminars were held regularly. Additionally, events in cooperation with the university hospital also took place, which I mostly attended together with lab members. By chance, my friend met a postdoc of the same nationality through a French

acquaintance in the same building, which allowed us to get to know a wider range of people through personal connections.

4 Everyday life and leisure

To be honest, my everyday life was not particularly unusual. At the beginning, I would go to the lab in the morning and have lunch with the lab members around noon. As time went on, I mostly spent my lunch breaks studying, and after lunch I would take a coffee break, chat with the lab members, and use that time to rest. It seemed that many PhD students and postdocs followed a similar routine.

One thing that really surprised me about Paris was the incredible variety and quality of the food. There were so many delicious options that I often found myself thinking about what to have for dinner even before leaving work. Even something as simple as bread was exceptionally good, and it became a small but important source of energy in my daily life. On weekends, I tried to visit as many different cafés as possible. Since I already enjoy café hopping, I preferred going to places frequented by locals rather than typical tourist spots. I usually visited major tourist attractions in the evenings on weekdays, when there were fewer people.

5 Financing

I received financial support through two main sources. The first was an Erasmus scholarship, and the second was a stipend. Fortunately, I was awarded an Erasmus+ grant agreement, which allowed me to cover approximately one and a half months of rent. At the same time, I also received a stipend from the host institute, amounting to around 1,000 euros per month. To the best of my knowledge, in France, if an internship lasts longer than two months and involves more than 32 working hours per week, even a voluntary internship is legally required to provide some form of financial compensation. I believe this information could be helpful for students who are planning to pursue an internship in France in the future.

6 Internship and studies

This internship was a valuable opportunity for me to gain experience in both wet-lab and dry-lab research. At the same time, I appreciated being able to directly experience the international culture of Paris. Although my master's program focuses primarily on biochemistry, I believe that informatics is an essential skill in modern science, which motivated me to take on this internship. Throughout the internship, I was able to seriously reflect on my future, including whether bioinformatics is truly a necessary field going forward, whether I have a genuine interest in it, whether I could pursue it as a long-term career, and whether I have the aptitude for this area. It also gave me the opportunity to think more concretely about my plans after completing my master's thesis. One clear conclusion I drew from this experience is that progress in the life sciences requires a well-balanced integration of wet-lab and

dry-lab approaches. I realized firsthand that science that leans too heavily toward only one side inevitably has limitations.

7 Conclusion

My internship at INSERM Paris was a highly valuable experience. I gained practical skills in both wet-lab experiments, including PCR, DNA extraction, and bacterial transformation, and dry-lab data analysis, such as SFB morphology quantification. The internship also taught me the importance of troubleshooting, careful planning, and data validation.

The flexible and supportive lab environment allowed frequent discussions with the PI and lab members, and participation in collaborative seminars and networking events helped me expand my professional and social contacts.

Overall, this internship gave me a deeper understanding of the integration of wet-lab and computational approaches in life sciences, strengthened my interest in pursuing research in this field, and provided insight into international scientific collaboration and daily lab life.