

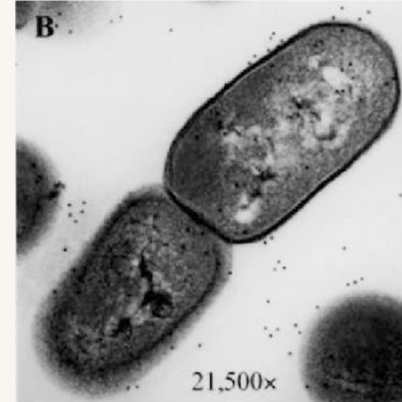
Estructura genética poblacional de aislados de *Listeria monocytogenes* y su asociación con la fuente de alimento

Casarin Perea Rodrigo; Cureño Sanchez Misael; Gonzalez

Servín Luis Israel; Juárez-García, G. Lizbeth

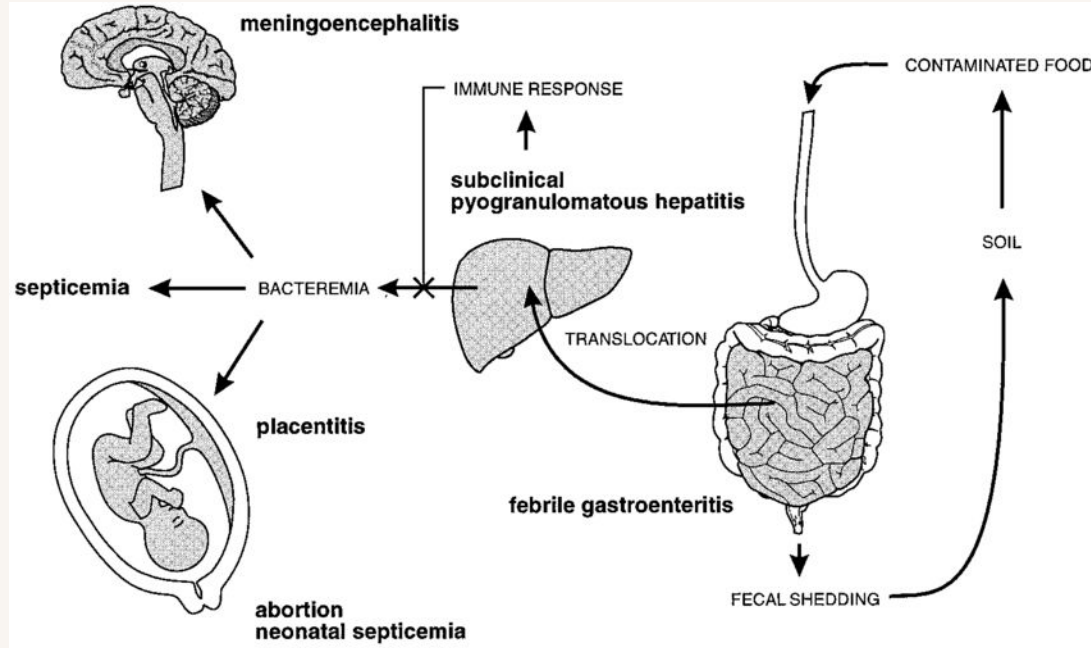
Introducción

- Patógenos transmitidos por los alimentos causan 600 millones de casos de enfermedades y 420 mil muertes al año.
- *L. monocytogenes* es una bacteria grampositiva, patógeno intracelular invasivo.
- 13 serotipos: 1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e y 7



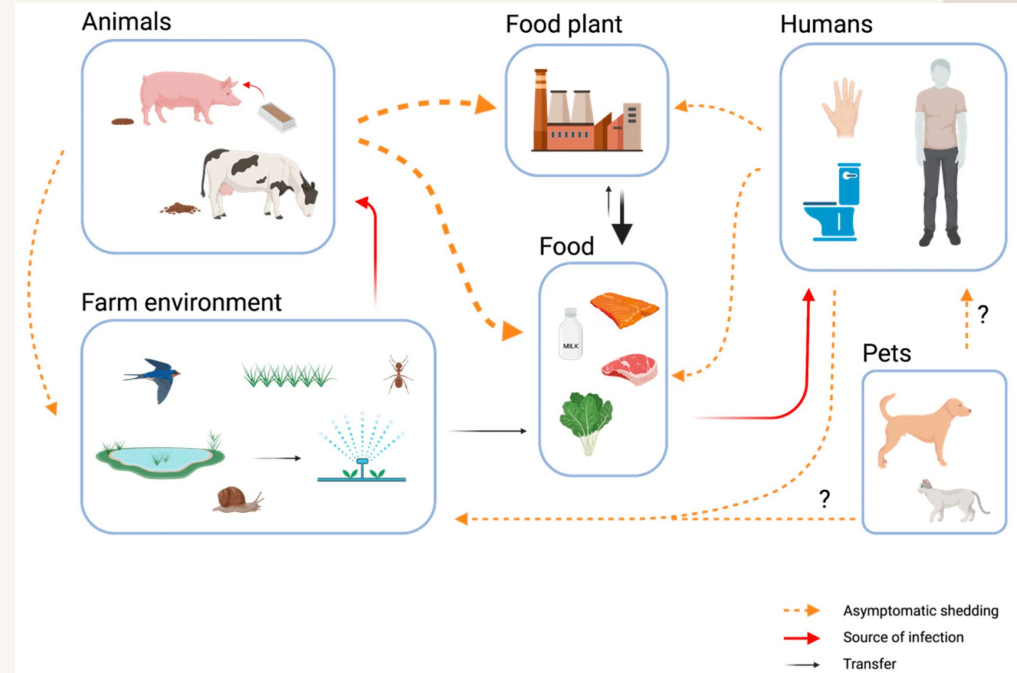
Listeria monocytogenes

- Listeriosis 20-30 % de mortalidad. Además, causar meningitis, gastroenteritis y septicemia como resultado de una infección.



Listeria monocytogenes

- Sobrevive a temperaturas extremas (1 a 45 °C) y concentraciones de sal.
- Forma biopelículas, le permite que contamine y persista en sustratos. Propagación rápida. Puede vivir durante años en superficies.
- Se ha encontrado en una variedad de alimentos.



Objetivos

- Conocer la estructura poblacional de aislados de *Listeria monocytogenes* de diversas fuentes de alimentos.
- Determinar si hay una relación entre poblaciones genéticamente distintas de *Listeria monocytogenes* y la fuente de alimento de la que proviene.

Hipótesis

- Al utilizar análisis de componentes principales para las secuencias de aislados de *L. monocytogenes* se determinará si los aislados pertenecen a una población homogénea o a una población que contiene subgrupos genéticamente distintos. Además, al comparar los subgrupos, se determinará si estos se distribuyen uniformemente en las fuentes de alimento.

Método

Datos

- Se usaron las secuencias de islados de L. monocytogenes 19 países.
- "queso", "carne cruda", "mariscos", "aguacates hass", "helado", "lechuga", "leche", "queso crema para untar", "papa", "carne cruda", "leche cruda", "verduras crudas", "salami", "camarones", "ambiente de matadero" y "salmón ahumado".
- Queso, otros lácteos, frutas/verduras, mariscos y carne.

Análisis de componentes

Supplementary Table S1. Metadata for the 504 <i>Listeria monocytogenes</i> isolates.						
	A	B	C	D	E	F
1	Supplementary Table S1. Metadata for the 504 <i>Listeria monocytogenes</i> isolates	GenBank assembly accession				
2						
3	Isolate Name	Food Source	Linear	Serotype	Population Assignment	
4	DAIRY_GCA_000681515_2_ASM68151v2_genomic	CHEESE	GCA_000681515.2	I	4b	2
5	DAIRY_GCA_000987705_1_ASM98770v1_genomic	DAIRY NON-CHEESE	GCA_000987705.1	I	1/2b	2
6	DAIRY_GCA_001317555_2_ASM131755v2_genomic	CHEESE	GCA_001317555.2	II	1/2c	10
7	DAIRY_GCA_001342195_2_ASM134219v2_genomic	CHEESE	GCA_001342195.2	I	4b	2
8	DAIRY_GCA_001463975_1_ASM146397v1_genomic	CHEESE	GCA_001463975.1	I	4b	2
9	DAIRY_GCA_001465135_2_ASM146513v2_genomic	CHEESE	GCA_001465135.2	II	1/2c	10
10	DAIRY_GCA_001466165_1_ASM146616v1_genomic	CHEESE	GCA_001466165.1	II	1/2a	6
11	DAIRY_GCA_001466195_1_ASM146619v1_genomic	CHEESE	GCA_001466195.1	II	1/2a	6
12	DAIRY_GCA_001548635_2_ASM154863v2_genomic	CHEESE	GCA_001548635.2	I	1/2b	1
13	DAIRY_GCA_002131935_1_ASM213193v1_genomic	CHEESE	GCA_002131935.1	I	4b	2
14	DAIRY_GCA_002132975_1_ASM213297v1_genomic	CHEESE	GCA_002132975.1	I	4b	2
15	DAIRY_GCA_002213825_1_ASM221382v1_genomic	CHEESE	GCA_002213825.1	II	1/2a	4
16	DAIRY_GCA_002213865_1_ASM221386v1_genomic	CHEESE	GCA_002213865.1	II	1/2a	4
17	DAIRY_GCA_002213945_1_ASM221394v1_genomic	CHEESE	GCA_002213945.1	I	4b	2
18	DAIRY_GCA_002213965_1_ASM221396v1_genomic	CHEESE	GCA_002213965.1	I	4b	2
19	DAIRY_GCA_002222785_1_ASM222278v1_genomic	CHEESE	GCA_002222785.1	I	4b	2
20	DAIRY_GCA_002445255_1_CESAN049216_genomic	DAIRY NON-CHEESE	GCA_002445255.1	II	1/2a	4


```
6 # Archivo excel del que se va a sacar el archivo .fasta
7 archivo_excel = 'Tabla.xlsx'
8 # Se lee el archivo Excel
9 df = pd.read_excel(archivo_excel)
10
11 # Nombre de la columna que contiene las secuencias de ADN
12 columna = 'GenBank assembly accession'
13
14 # Se crean objetos SeqRecord para cada secuencia y se guardan en una lista
15 obj = []
16 for index, row in df.iterrows():
17     sequence = str(row[columna])
18     if pd.notna(sequence):
19         record = SeqRecord(Seq(sequence), id=str(index), description="")
20         obj.append(record)
21 # Archivo .fasta resultante
22 archivo_fasta = 'salida.fasta'
23
24 # Se guardan los registros en un archivo .fasta
25 with open(archivo_fasta, "w") as fasta_file:
26     for secuencia in obj:
27         if pd.notna(secuencia.seq):
28             fasta_file.write(str(secuencia.seq) + "\n")
29
30 print(f"Archivo .fasta creado con éxito.")
```

```

if not os.path.exists(GENOMES_FOLDER):
    os.makedirs(GENOMES_FOLDER)

with open('salida.fasta') as f:
    COMANDO = 'mafft --auto GCA_000681515.2_ASM68151v2_genomic.fna > first_align.fasta'
    SEGUNDO_COMANDO = f'mafft --6merpair --thread -1 --keeplength --addfragments {others} first_align.fasta > alineamiento.fasta'

# Si se está ejecutando desde Windows concatenamos al principio el comando wsl
if platform.system() == 'Windows':
    PRIMER_COMANDO = 'wsl.exe ' + PRIMER_COMANDO
    SEGUNDO_COMANDO = 'wsl.exe ' + SEGUNDO_COMANDO

os.system(PRIMER_COMANDO)
os.system(SEGUNDO_COMANDO)

accessions = [x.strip() for x in accessions]
accessions = filter(lambda x: x != '' and x != 'nan', accessions)

for accession in accessions:
    url = f'https://api.ncbi.nlm.nih.gov/datasets/v2alpha/genome/accession/{accession}/download?include_annotation_type=GENOME_FASTA'
    r = requests.get(url)

    with open(f'{GENOMES_FOLDER}/{accession}.zip', 'wb') as fasta:
        fasta.write(r.content)
    print(f'Secuencia {accession} descargada con éxito.')
    print(f'Extrayendo el zip')

```

```
# Creamos un nuevo directorio llamado data
if not os.path.exists('data'):
    os.makedirs('data')

# We get the absolute path of the current directory
data_dir = os.path.join(os.getcwd(), 'data')

# Copiamos todos los archivos dentro de la carpeta secuencia que están
# en formato .fna a la carpeta data de manera recursiva, esto es,
# incluyendo los archivos en los subdirectorios de secuencias.

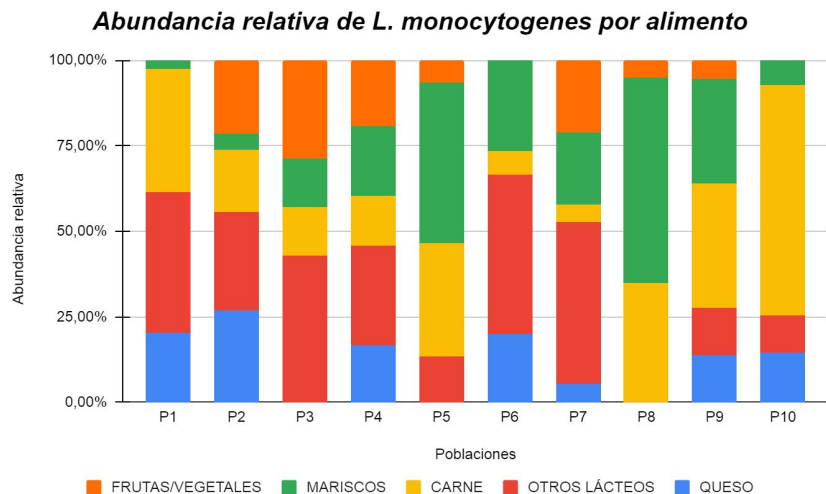
for root, dirs, files in os.walk(GENOMES_FOLDER):
    for file in files:
        if file.endswith('.fna'):
            print(f'Copiando {file} a la carpeta {data_dir}')
            # Lo copiamos a la carpeta data
            shutil.copyfile(os.path.join(root, file), os.path.join(data_dir, file))
```

Resultados y discusión

Pearson's Chi-squared test

```
data: datos1
```

```
X-squared = 190.16, df = 36, p-value < 2.2e-16
```



Los aislados de las 10 poblaciones no se distribuyen uniformemente entre las fuentes de alimentos de las que provienen.

Conclusión

- De las 10 poblaciones principales de *Listeria monocytogenes*, se encontró que la población 2 predomina en Quesos y Frutas/Verduras, las poblaciones 5, 8 y 9 predomina en Mariscos, y finalmente, la población 10 predominó en Carne. Esto indicó que existen asociaciones con alimentos particulares probablemente por sus diferencias genéticas.

Referencias

- Abraham, G. y Inouye, M. (2014). Fast principal component analysis of large-scale genome-wide data. PloS one, 9(4), e93766.
- Bechtel, T. y Gibbons, J. (2021). Population genomic analysis of *Listeria monocytogenes* from food reveals substrate-specific genome variation. Frontiers in Microbiology, 12, 620033.
- Bintsis, T. (2017). Foodborne pathogens. AIMS microbiology, 3(3), 529.
- Sibanda, T. y Buys, E. (2022). *Listeria monocytogenes* pathogenesis: The role of stress adaptation. Microorganisms, 10(8), 1522.

■