**SUPPLEMENTARY FILE**

**Supplementary Data 1:**

# -\*- coding: utf-8 -\*-

**import** requests

**import** pandas **as** pd

**from** time **import** sleep

**import** os

**def** **get\_paper\_count**(term, year, retries=3):

base\_url = "https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi"

params = {

"db": "pubmed",

"term": term,

"mindate": f"{year}/01/01",

"maxdate": f"{year}/12/31",

"datetype": "pdat",

"retmode": "json"

}

**for** attempt **in** range(retries):

**try**:

response = requests.get(base\_url, params=params)

response.raise\_for\_status() # Check for HTTP errors

data = response.json()

count = int(data["esearchresult"]["count"])

**return** count

**except** requests.exceptions.RequestException **as** e:

print(f"HTTP error on attempt {attempt + 1}: {e}")

sleep(2) # Wait for 2 seconds before retrying

**except** ValueError:

print(f"Error parsing JSON response for year {year} and term '{term}' on attempt {attempt + 1}")

sleep(2) # Wait for 2 seconds before retrying

**return** None

# Define the search terms for each molecular marker

markers = ["RFLP", "RAPD", "ISSR", "AFLP", "SSR", "EST-SSR", "SNP", "GBS"]

base\_search\_term = "(genetic diversity) AND (endangered)"

# Initialize list to store results

results = []

# Loop over each marker and year

**for** marker **in** markers:

search\_term = f"{base\_search\_term} ({marker})"

**for** year **in** range(1979, 2024):

print(f"Fetching data for marker '{marker}' in the year {year}...")

paper\_count = get\_paper\_count(search\_term, year)

**if** paper\_count **is** **not** None:

results.append({

"Year": year,

"Marker": marker,

"Number of Papers": paper\_count

})

#NCBI rate limit 2/sec

sleep(1.0)

# Convert the results list to a DataFrame

df = pd.DataFrame(results)

# Display the DataFrame

print(df)

#save the CSV file

output\_dir = "path/to/directory"

output\_file = os.path.join(output\_dir, "ncbi\_genetic\_diversity\_endangered\_markers.csv")

# Save the DataFrame to a CSV file in the specified directory

df.to\_csv(output\_file, index=False)

# -\*- coding: utf-8 -\*-

**import** pandas **as** pd

**import** matplotlib.pyplot **as** plt

**import** seaborn **as** sns

# Load the data

file\_path = 'path/to/file'

df = pd.read\_csv(file\_path)

# Display the first few rows

print(df.head())

# Create bar plot

plt.figure(figsize=(14, 7))

sns.barplot(data=df, x='Year', y='Number of Papers', hue='Marker')

plt.title('Assessment of Genetic Diversity Using Different Molecular Markers (1979-2023)')

plt.xlabel('Year')

plt.ylabel('Number of Papers present at NCBI Database')

plt.legend(title='Molecular Markers Used')

plt.xticks(rotation=90)

plt.grid(True)

#save and show plot

plt.savefig('plot.png')

plt.show()

**Supplementary Data 2:**

|  |  |  |
| --- | --- | --- |
| **Year** | **Marker** | **Number of Papers** |
| 1979 | RFLP | 0 |
| 1980 | RFLP | 0 |
| 1981 | RFLP | 0 |
| 1982 | RFLP | 0 |
| 1983 | RFLP | 0 |
| 1984 | RFLP | 0 |
| 1985 | RFLP | 0 |
| 1986 | RFLP | 0 |
| 1987 | RFLP | 0 |
| 1988 | RFLP | 1 |
| 1989 | RFLP | 0 |
| 1990 | RFLP | 0 |
| 1991 | RFLP | 1 |
| 1992 | RFLP | 1 |
| 1993 | RFLP | 2 |
| 1994 | RFLP | 1 |
| 1995 | RFLP | 2 |
| 1996 | RFLP | 1 |
| 1997 | RFLP | 1 |
| 1998 | RFLP | 0 |
| 1999 | RFLP | 1 |
| 2000 | RFLP | 1 |
| 2001 | RFLP | 3 |
| 2002 | RFLP | 1 |
| 2003 | RFLP | 3 |
| 2004 | RFLP | 5 |
| 2005 | RFLP | 3 |
| 2006 | RFLP | 3 |
| 2007 | RFLP | 0 |
| 2008 | RFLP | 4 |
| 2009 | RFLP | 4 |
| 2010 | RFLP | 5 |
| 2011 | RFLP | 3 |
| 2012 | RFLP | 4 |
| 2013 | RFLP | 4 |
| 2014 | RFLP | 4 |
| 2015 | RFLP | 3 |
| 2016 | RFLP | 2 |
| 2017 | RFLP | 0 |
| 2018 | RFLP | 1 |
| 2019 | RFLP | 3 |
| 2020 | RFLP | 0 |
| 2021 | RFLP | 1 |
| 2022 | RFLP | 1 |
| 2023 | RFLP | 0 |
| 1979 | RAPD | 0 |
| 1980 | RAPD | 0 |
| 1981 | RAPD | 0 |
| 1982 | RAPD | 0 |
| 1983 | RAPD | 0 |
| 1984 | RAPD | 0 |
| 1985 | RAPD | 0 |
| 1986 | RAPD | 0 |
| 1987 | RAPD | 0 |
| 1988 | RAPD | 0 |
| 1989 | RAPD | 0 |
| 1990 | RAPD | 0 |
| 1991 | RAPD | 0 |
| 1992 | RAPD | 0 |
| 1993 | RAPD | 0 |
| 1994 | RAPD | 2 |
| 1995 | RAPD | 2 |
| 1996 | RAPD | 2 |
| 1997 | RAPD | 3 |
| 1998 | RAPD | 2 |
| 1999 | RAPD | 1 |
| 2000 | RAPD | 2 |
| 2001 | RAPD | 5 |
| 2002 | RAPD | 4 |
| 2003 | RAPD | 3 |
| 2004 | RAPD | 4 |
| 2005 | RAPD | 5 |
| 2006 | RAPD | 3 |
| 2007 | RAPD | 6 |
| 2008 | RAPD | 2 |
| 2009 | RAPD | 4 |
| 2010 | RAPD | 5 |
| 2011 | RAPD | 7 |
| 2012 | RAPD | 5 |
| 2013 | RAPD | 5 |
| 2014 | RAPD | 4 |
| 2015 | RAPD | 2 |
| 2016 | RAPD | 2 |
| 2017 | RAPD | 1 |
| 2018 | RAPD | 1 |
| 2019 | RAPD | 2 |
| 2020 | RAPD | 0 |
| 2021 | RAPD | 3 |
| 2022 | RAPD | 0 |
| 2023 | RAPD | 0 |
| 1979 | ISSR | 0 |
| 1980 | ISSR | 0 |
| 1981 | ISSR | 0 |
| 1982 | ISSR | 0 |
| 1983 | ISSR | 0 |
| 1984 | ISSR | 0 |
| 1985 | ISSR | 0 |
| 1986 | ISSR | 0 |
| 1987 | ISSR | 0 |
| 1988 | ISSR | 0 |
| 1989 | ISSR | 0 |
| 1990 | ISSR | 0 |
| 1991 | ISSR | 0 |
| 1992 | ISSR | 0 |
| 1993 | ISSR | 0 |
| 1994 | ISSR | 0 |
| 1995 | ISSR | 0 |
| 1996 | ISSR | 0 |
| 1997 | ISSR | 0 |
| 1998 | ISSR | 0 |
| 1999 | ISSR | 0 |
| 2000 | ISSR | 0 |
| 2001 | ISSR | 0 |
| 2002 | ISSR | 0 |
| 2003 | ISSR | 0 |
| 2004 | ISSR | 1 |
| 2005 | ISSR | 3 |
| 2006 | ISSR | 4 |
| 2007 | ISSR | 3 |
| 2008 | ISSR | 5 |
| 2009 | ISSR | 6 |
| 2010 | ISSR | 2 |
| 2011 | ISSR | 5 |
| 2012 | ISSR | 6 |
| 2013 | ISSR | 8 |
| 2014 | ISSR | 8 |
| 2015 | ISSR | 10 |
| 2016 | ISSR | 4 |
| 2017 | ISSR | 1 |
| 2018 | ISSR | 3 |
| 2019 | ISSR | 3 |
| 2020 | ISSR | 5 |
| 2021 | ISSR | 6 |
| 2022 | ISSR | 3 |
| 2023 | ISSR | 4 |
| 1979 | AFLP | 0 |
| 1980 | AFLP | 0 |
| 1981 | AFLP | 0 |
| 1982 | AFLP | 0 |
| 1983 | AFLP | 0 |
| 1984 | AFLP | 0 |
| 1985 | AFLP | 0 |
| 1986 | AFLP | 0 |
| 1987 | AFLP | 0 |
| 1988 | AFLP | 0 |
| 1989 | AFLP | 0 |
| 1990 | AFLP | 0 |
| 1991 | AFLP | 0 |
| 1992 | AFLP | 0 |
| 1993 | AFLP | 0 |
| 1994 | AFLP | 0 |
| 1995 | AFLP | 0 |
| 1996 | AFLP | 1 |
| 1997 | AFLP | 0 |
| 1998 | AFLP | 0 |
| 1999 | AFLP | 0 |
| 2000 | AFLP | 2 |
| 2001 | AFLP | 1 |
| 2002 | AFLP | 0 |
| 2003 | AFLP | 3 |
| 2004 | AFLP | 3 |
| 2005 | AFLP | 1 |
| 2006 | AFLP | 8 |
| 2007 | AFLP | 3 |
| 2008 | AFLP | 4 |
| 2009 | AFLP | 3 |
| 2010 | AFLP | 4 |
| 2011 | AFLP | 13 |
| 2012 | AFLP | 9 |
| 2013 | AFLP | 6 |
| 2014 | AFLP | 8 |
| 2015 | AFLP | 4 |
| 2016 | AFLP | 5 |
| 2017 | AFLP | 6 |
| 2018 | AFLP | 4 |
| 2019 | AFLP | 0 |
| 2020 | AFLP | 1 |
| 2021 | AFLP | 2 |
| 2022 | AFLP | 3 |
| 2023 | AFLP | 3 |
| 1979 | SSR | 0 |
| 1980 | SSR | 0 |
| 1981 | SSR | 0 |
| 1982 | SSR | 0 |
| 1983 | SSR | 0 |
| 1984 | SSR | 0 |
| 1985 | SSR | 0 |
| 1986 | SSR | 0 |
| 1987 | SSR | 0 |
| 1988 | SSR | 0 |
| 1989 | SSR | 0 |
| 1990 | SSR | 0 |
| 1991 | SSR | 0 |
| 1992 | SSR | 0 |
| 1993 | SSR | 0 |
| 1994 | SSR | 0 |
| 1995 | SSR | 0 |
| 1996 | SSR | 0 |
| 1997 | SSR | 0 |
| 1998 | SSR | 0 |
| 1999 | SSR | 0 |
| 2000 | SSR | 0 |
| 2001 | SSR | 0 |
| 2002 | SSR | 0 |
| 2003 | SSR | 2 |
| 2004 | SSR | 1 |
| 2005 | SSR | 2 |
| 2006 | SSR | 1 |
| 2007 | SSR | 3 |
| 2008 | SSR | 2 |
| 2009 | SSR | 2 |
| 2010 | SSR | 2 |
| 2011 | SSR | 5 |
| 2012 | SSR | 8 |
| 2013 | SSR | 8 |
| 2014 | SSR | 8 |
| 2015 | SSR | 11 |
| 2016 | SSR | 15 |
| 2017 | SSR | 20 |
| 2018 | SSR | 15 |
| 2019 | SSR | 13 |
| 2020 | SSR | 20 |
| 2021 | SSR | 7 |
| 2022 | SSR | 4 |
| 2023 | SSR | 9 |
| 1979 | EST-SSR | 0 |
| 1980 | EST-SSR | 0 |
| 1981 | EST-SSR | 0 |
| 1982 | EST-SSR | 0 |
| 1983 | EST-SSR | 0 |
| 1984 | EST-SSR | 0 |
| 1985 | EST-SSR | 0 |
| 1986 | EST-SSR | 0 |
| 1987 | EST-SSR | 0 |
| 1988 | EST-SSR | 0 |
| 1989 | EST-SSR | 0 |
| 1990 | EST-SSR | 0 |
| 1991 | EST-SSR | 0 |
| 1992 | EST-SSR | 0 |
| 1993 | EST-SSR | 0 |
| 1994 | EST-SSR | 0 |
| 1995 | EST-SSR | 0 |
| 1996 | EST-SSR | 0 |
| 1997 | EST-SSR | 0 |
| 1998 | EST-SSR | 0 |
| 1999 | EST-SSR | 0 |
| 2000 | EST-SSR | 0 |
| 2001 | EST-SSR | 0 |
| 2002 | EST-SSR | 0 |
| 2003 | EST-SSR | 0 |
| 2004 | EST-SSR | 0 |
| 2005 | EST-SSR | 0 |
| 2006 | EST-SSR | 0 |
| 2007 | EST-SSR | 0 |
| 2008 | EST-SSR | 0 |
| 2009 | EST-SSR | 0 |
| 2010 | EST-SSR | 1 |
| 2011 | EST-SSR | 0 |
| 2012 | EST-SSR | 1 |
| 2013 | EST-SSR | 1 |
| 2014 | EST-SSR | 0 |
| 2015 | EST-SSR | 1 |
| 2016 | EST-SSR | 4 |
| 2017 | EST-SSR | 3 |
| 2018 | EST-SSR | 3 |
| 2019 | EST-SSR | 3 |
| 2020 | EST-SSR | 5 |
| 2021 | EST-SSR | 0 |
| 2022 | EST-SSR | 1 |
| 2023 | EST-SSR | 1 |
| 1979 | SNP | 0 |
| 1980 | SNP | 0 |
| 1981 | SNP | 0 |
| 1982 | SNP | 0 |
| 1983 | SNP | 0 |
| 1984 | SNP | 0 |
| 1985 | SNP | 0 |
| 1986 | SNP | 0 |
| 1987 | SNP | 0 |
| 1988 | SNP | 0 |
| 1989 | SNP | 0 |
| 1990 | SNP | 0 |
| 1991 | SNP | 0 |
| 1992 | SNP | 0 |
| 1993 | SNP | 0 |
| 1994 | SNP | 0 |
| 1995 | SNP | 0 |
| 1996 | SNP | 0 |
| 1997 | SNP | 0 |
| 1998 | SNP | 0 |
| 1999 | SNP | 0 |
| 2000 | SNP | 0 |
| 2001 | SNP | 0 |
| 2002 | SNP | 0 |
| 2003 | SNP | 0 |
| 2004 | SNP | 0 |
| 2005 | SNP | 0 |
| 2006 | SNP | 1 |
| 2007 | SNP | 0 |
| 2008 | SNP | 1 |
| 2009 | SNP | 2 |
| 2010 | SNP | 1 |
| 2011 | SNP | 6 |
| 2012 | SNP | 7 |
| 2013 | SNP | 6 |
| 2014 | SNP | 7 |
| 2015 | SNP | 9 |
| 2016 | SNP | 15 |
| 2017 | SNP | 22 |
| 2018 | SNP | 17 |
| 2019 | SNP | 20 |
| 2020 | SNP | 12 |
| 2021 | SNP | 23 |
| 2022 | SNP | 27 |
| 2023 | SNP | 21 |
| 1979 | GBS | 0 |
| 1980 | GBS | 0 |
| 1981 | GBS | 0 |
| 1982 | GBS | 0 |
| 1983 | GBS | 0 |
| 1984 | GBS | 0 |
| 1985 | GBS | 0 |
| 1986 | GBS | 0 |
| 1987 | GBS | 0 |
| 1988 | GBS | 0 |
| 1989 | GBS | 0 |
| 1990 | GBS | 0 |
| 1991 | GBS | 0 |
| 1992 | GBS | 0 |
| 1993 | GBS | 0 |
| 1994 | GBS | 0 |
| 1995 | GBS | 0 |
| 1996 | GBS | 0 |
| 1997 | GBS | 0 |
| 1998 | GBS | 0 |
| 1999 | GBS | 0 |
| 2000 | GBS | 0 |
| 2001 | GBS | 0 |
| 2002 | GBS | 0 |
| 2003 | GBS | 0 |
| 2004 | GBS | 0 |
| 2005 | GBS | 0 |
| 2006 | GBS | 0 |
| 2007 | GBS | 0 |
| 2008 | GBS | 0 |
| 2009 | GBS | 0 |
| 2010 | GBS | 0 |
| 2011 | GBS | 0 |
| 2012 | GBS | 0 |
| 2013 | GBS | 0 |
| 2014 | GBS | 0 |
| 2015 | GBS | 0 |
| 2016 | GBS | 1 |
| 2017 | GBS | 2 |
| 2018 | GBS | 4 |
| 2019 | GBS | 1 |
| 2020 | GBS | 0 |
| 2021 | GBS | 2 |
| 2022 | GBS | 2 |
| 2023 | GBS | 6 |