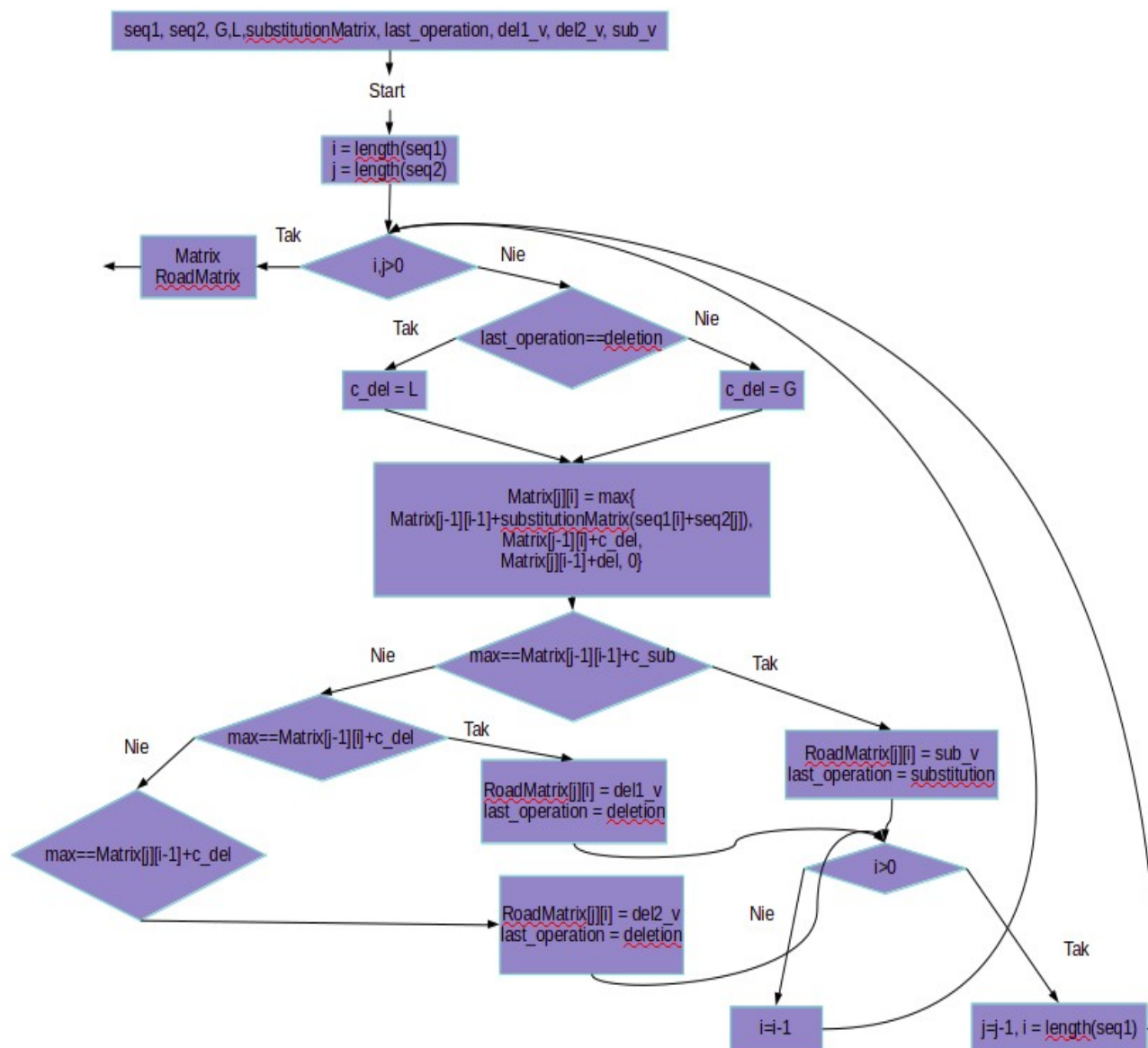


## 1. Schemat blokowy

### 1.1 Tworzenie macierzy drogi oraz macierzy punktacji:

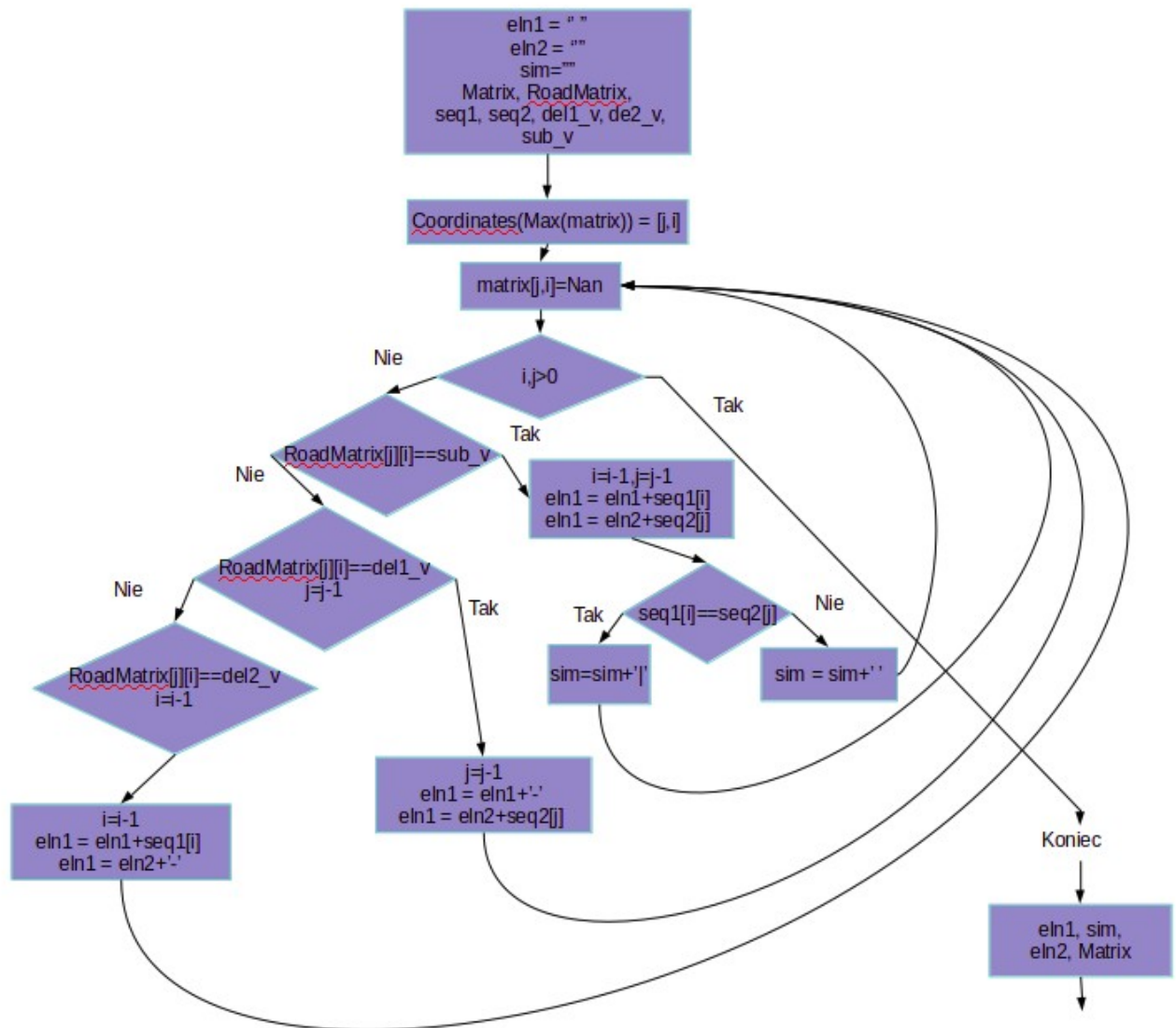


del1\_v – wartość w macierzy drogi świadcząca o tym, iż doszło do delecji (przejście z góry)

del2\_v – wartość w macierzy drogi świadcząca o tym, iż doszło do delecji (przejście z lewej strony)

sub\_v – wartość w macierzy drogi świadcząca o tym, iż doszło do substytucji (przejście na skos)

1.2 Wyznaczenie i oznaczenie (za pomocą wartości Nan) optymalnej ścieżki dopasowania:



del1\_v – wartość w macierzy drogi świadcząca o tym, iż doszło do delecji (przejście z góry)

del2\_v – wartość w macierzy drogi świadcząca o tym, iż doszło do delecji (przejście z lewej strony)

sub\_v – wartość w macierzy drogi świadcząca o tym, iż doszło do substytucji (przejście na skos)

## 2. Analiza złożoności obliczeniowej czasowej i pamięciowej

2.1 Tworzenie macierzy drogi oraz macierzy punktacji:

- **czasowa:**  $O(m \cdot n)$
- **pamięciowa:**  $O(m \cdot n)$

Gdzie m i n to odpowiednio długości pierwszej i drugiej sekwencji

2.2 Wyznaczenie i oznaczenie (za pomocą wartości Nan) optymalnej ścieżki dopasowania:

- **czasowa:**  $O(m+n)$
- **pamięciowa:**  $(m+n)$

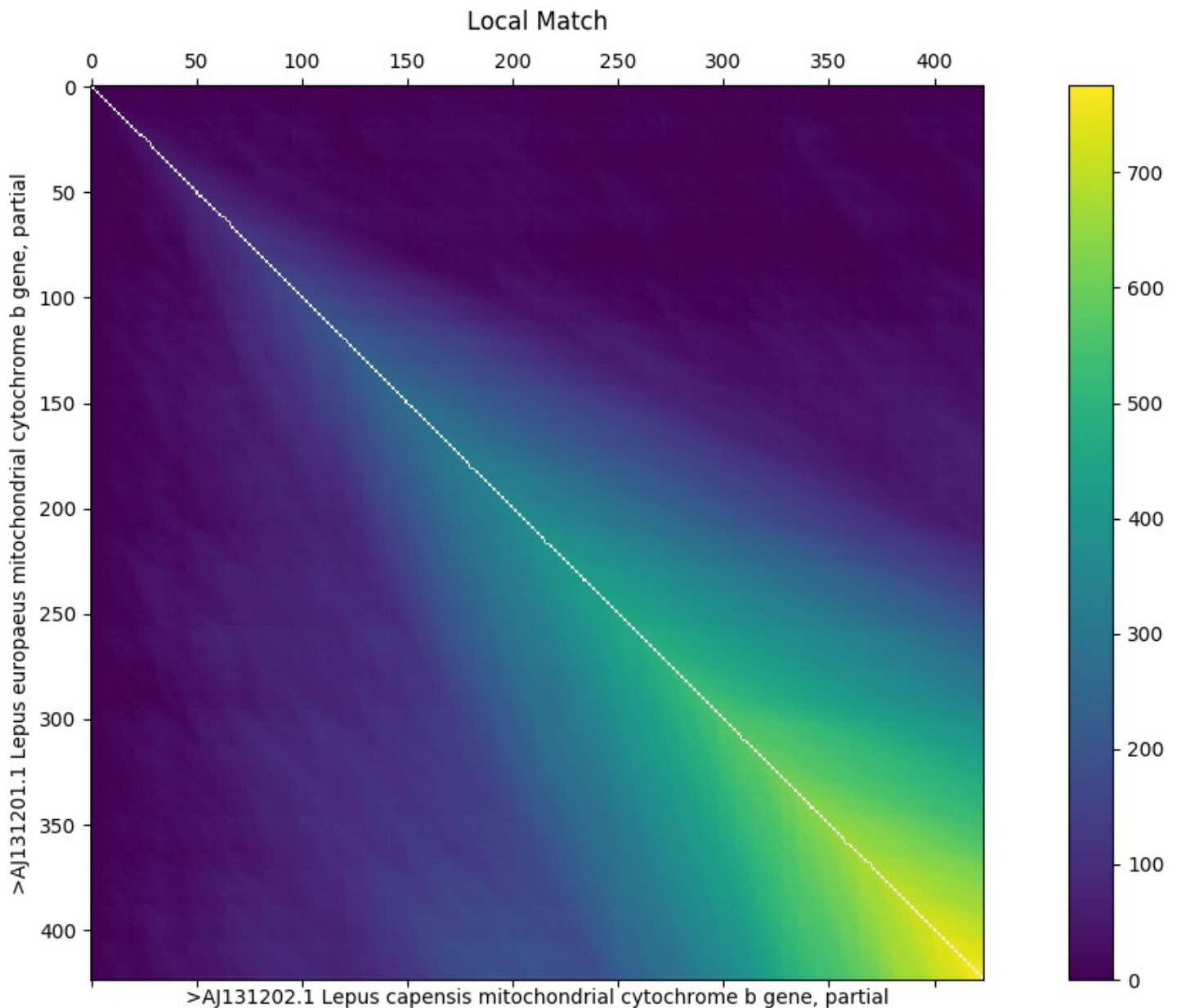
Gdzie m i n to odpowiednio długości pierwszej i drugiej sekwencji

### 2.3 Całokształt:

- **czasowa:**  $O(m \cdot n)$
- **pamięciowa:**  $O(m \cdot n)$

## 3. Porównanie przykładowych par sekwencji ewolucyjnie powiązanych i niepowiązanych

### 3.1 Powiązanych:



Rys 1, Dopasowanie globalne dla genów kodujących cytochrom b dla zająca szara i zająca pustynnego

Sekwencje zostały wybrane nie ze względu na duże pokrewieństwo danych gatunków, lecz ze względu na to, że są tych samej długości.

```
ATGACCAACATTTCGTAACGACATCCCC-TACTAAAAATTGTTAACCACCTC-CCTAATCGACCTTCCC-GCCCCCTCAAACATCTCTGCCTGATGAACTTCGGCTCTCTATTGGGATTATGCCTAATAATCCAGATCCTAACTGGCTTATT-  
|||||  
ATGACCAACATTTCGTAACGCA-CCCCCTACTAAAAATTGTTAACCACCTCTC-TAATCGACCT-CCCCGCCCCCTCAAACATCTCTGCCTGATGAACTTCGGCTCTCTATTGGGATTATGCCTAATAATCCAGATCCTAACTGGCTTATT
```

CCATAGCCATACACTACACATCAGACACAGC-AACAGCATTCTCCTCAGTTACACATATTTGCCGAGA-CGTAAACTACGGCTGACT-TATTGTTACTTACACGCTAATGGAGCATCAATATTTCTTTATTTGCTTATATATACATGTAGGCCGTGGAATCTA  
|||||  
C-TAGCCATACACTACACATCAGACACAGCTA-CAGCATTCTCCTCAGTTACACATATTTGCCGAGATG-TAAACTACGGCTGACTC-ATTGTTACTTACACGCTAATGGAGCATCAATATTTCTTTATTTGCTTATATATACATGTAGGCCGTGGAATCTA  
|||||  
-TTACGGCTCATATACTTACCTAGAAACCTGAAA-TATTGGCATTATTCTA-TTATTCGCAGTTA-TGGCTACAGCATTATAGGCTA-TGTCCTCCCATGAGGACAAATATCATTCTGA  
|||||  
CT-ACGGCTCATATACTTACCTAGAAACCTGAAAC-ATTGGCATTATTCTACT-ATTCGCAGT-AATGGCTACAGCATTATAGGCTAC-GTCCTCCCATGAGGACAAATATCATTCTGA

### Fragmenty statystyk:

```
# Mode: similarity
```

```
# Punctuation: {'AA': 2, 'CA': -7, 'GA': -5, 'TA': -7, 'AC': -7, 'CC': 2, 'GC': 7, 'TC': -5, 'AG': -5, 'CG': -7, 'GG': 2, 'TG': -7, 'AT': -7, 'CT': -5, 'GT': -7, 'TT': 2}
```

```
# Gap: -2
```

# G: -15

# L: -2

# Score: 779.0

```
# Length: 435
```

```
# Identity: 410/435 (94%)
```

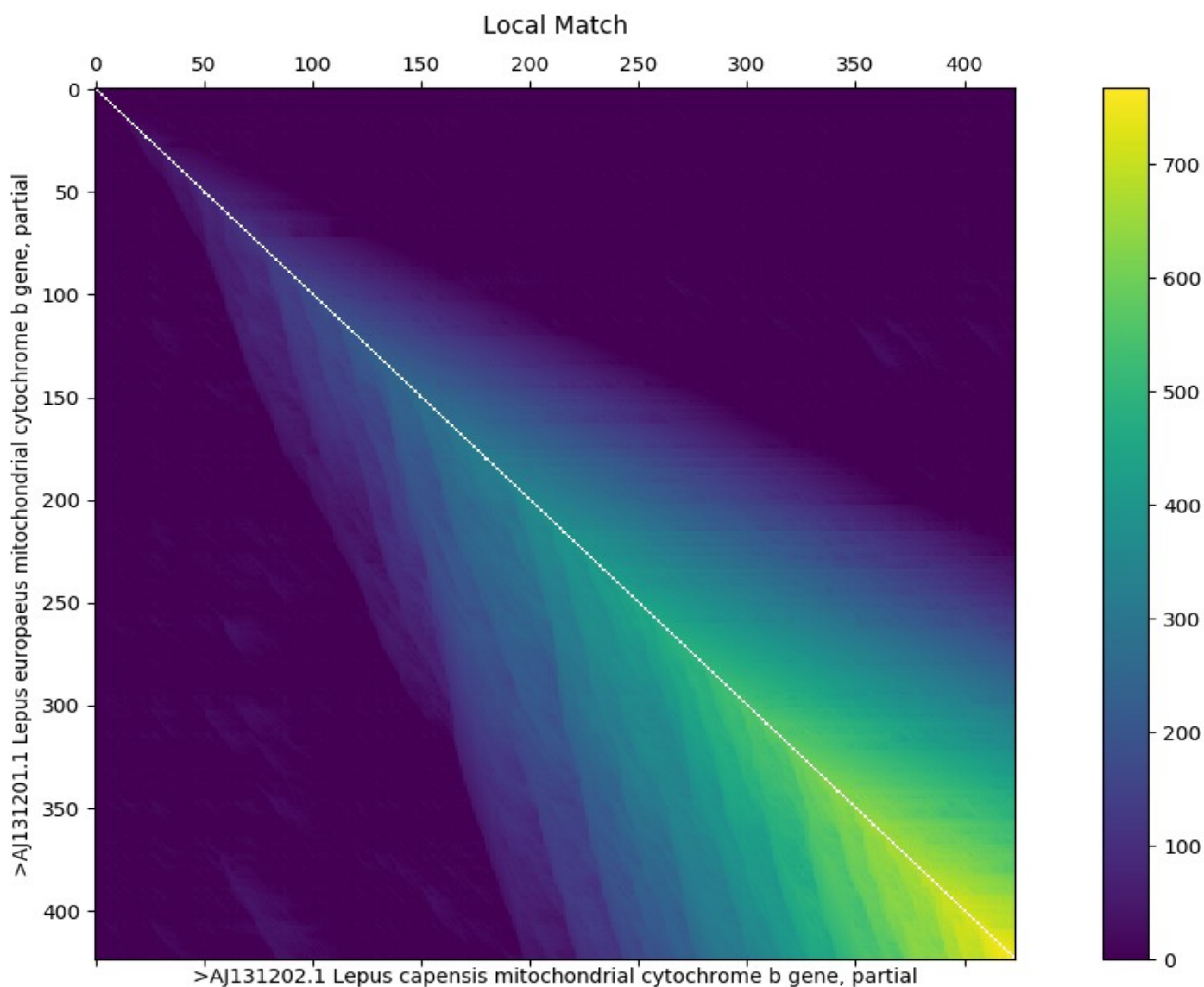
# Gaps: 24/435 (5%)

### Fasta Format:

```
>seq1 1-423
```

```
>seq2 1-423
```

Dopasowanie za pomocą algorytmu rozróżniającego karę rozpoczęcia oraz kontynuacji przerwy:



Rys 2, Dopasowanie globalne dla genów kodujących cytochrom b dla zająca szara i zająca pustynnego za pomocą algorytmu rozróżniającego karę rozpoczęcia oraz kontynuacji przerwy

### Fragmenty statystyk:

#### # Mode: similarity

```
# Punctuation: {'AA': 2, 'CA': -7, 'GA': -5, 'TA': -7, 'AC': -7, 'CC': 2, 'GC': 7, 'TC': -5, 'AG': -5, 'CG': -7, 'GG': 2, 'TG': -7, 'AT': -7, 'CT': -5, 'GT': -7, 'TT': 2}
# Gap: -2
# G: -15
# L: -2
# Score: 771.0
# Length: 434
# Identity: 411/434 (94%)
# Gaps: 22/434 (5%)
```

#### Fasta Format:

```
>seq1 1-423
```

```
>seq2 1-423
```

ATGACCAACATTCGTAAACGCA-TCCCCTACTAAAAATTGTTAACCACTC-CCTAATCGACCT-TCCCGCCCCCTCAAACATCTCTGCCTGATGAACTTCGGCTCTCTATTGGGATTATGCCTAATAATCCAGATCCTAACTGGCTTATT-  
|||||  
ATGACCAACATTCGTAAACGCAC-CCCCTACTAAAAATTGTTAACCACTCT-CTAATCGACCTC-CCCGCCCCCTCAAACATCTCTGCCTGATGAACTTCGGCTCTCTATTGGGATTATGCCTAATAATCCAGATCCTAACTGGCTTATT-  
|||||  
CCTAGCCATACACTACACATCAGACACAGC-AACAGCATTCTCCTCAGTTACACATATTTGCCGAGA-CGTAAACTACGGCTGACT-TATTCGTTACTTACACGCTAATGGAGCATCAATATTCTTTATTTGCTTATATACATGTAGGCCGTGGAATCTA-  
|||||  
-CTAGCCATACACTACACATCAGACACAGCT-ACAGCATTCTCCTCAGTTACACATATTTGCCGAGAT-GTAAACTACGGCTGACTC-ATTCGTTACTTACACGCTAATGGAGCATCAATATTCTTTATTTGCTTATATACATGTAGGCCGTGGAATCTAC-  
|||||  
TTACGGCTCATATACTTACCTAGAAACCTGAAA-TATTGGCATTATTCTA-TTATTCGCAGTTATGGCTACAGCATTTCATAGGCTA-TGTCCTCCCATGAGGACAAATATCATTCTGA-  
|||||  
-TACGGCTCATATACTTACCTAGAAACCTGAAAC-ATTGGCATTATTCTAC-TATTCGCAGTAATGGCTACAGCATTTCATAGGCTAC-GTCCTCCCATGAGGACAAATATCATTCTGA-  
|||||

### 3.2 Niepowiązanych

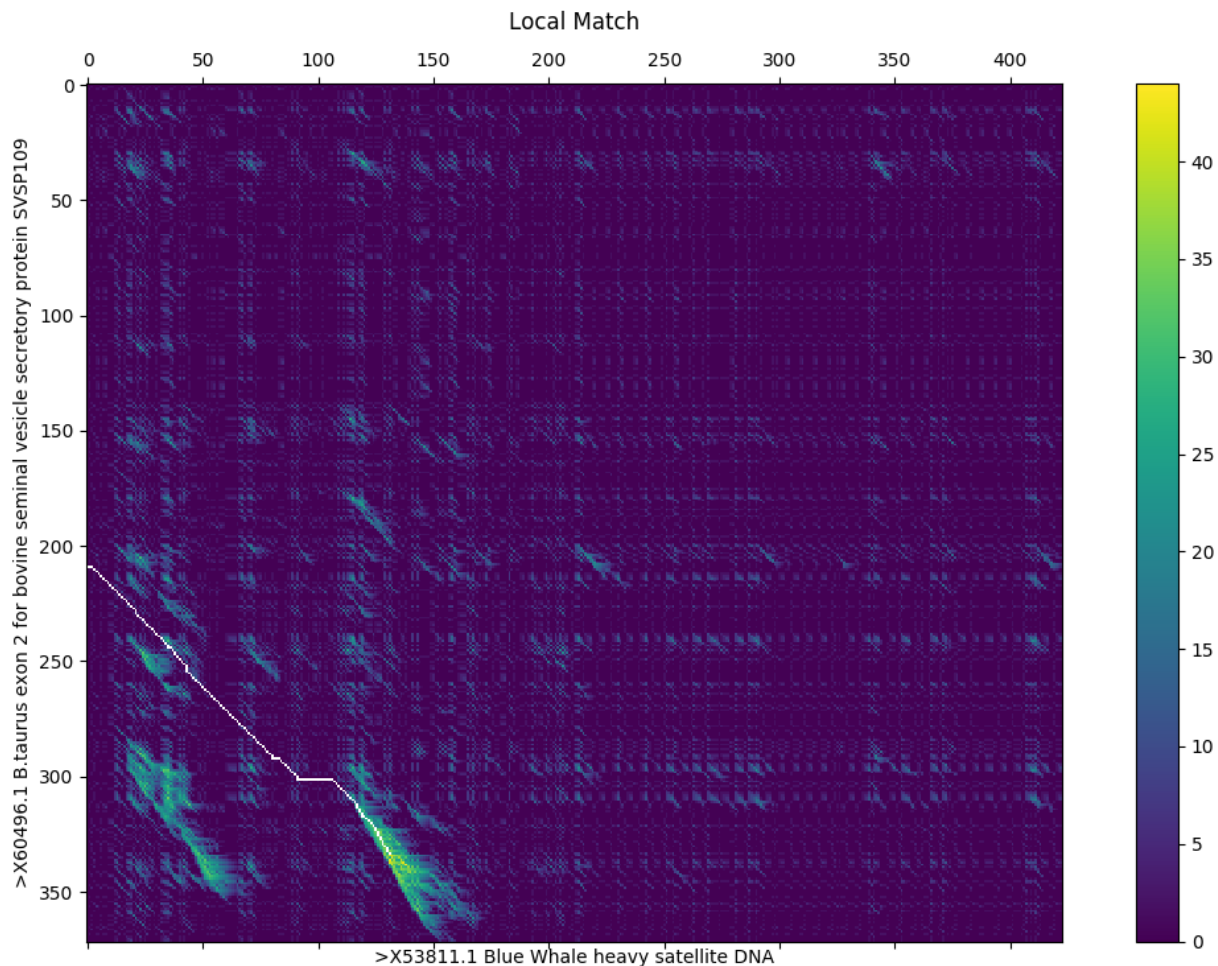








Dopasowanie za pomocą algorytmu rozróżniającego karę rozpoczęcia oraz kontynuacji przerwy:



Rys 4. Dopasowanie lokalne dla przypadkowo wybranych genów niepowiązanych ewolucyjnie ze sobą za pomocą algorytmu rozróżniającego karę rozpoczęcia oraz kontynuacji przerwy

Fragment statystyk:

```
# Mode: similarity
```

```
# Punctuation: {'AA': 2, 'CA': -7, 'GA': -5, 'TA': -7, 'AC': -7, 'CC': 2, 'GC': 7, 'TC': -5, 'AG': -5, 'CG': -7, 'GG': 2, 'TG': -7, 'AT': -7, 'CT': -5, 'GT': -7, 'TT': 2}
# Gap: -2
# G: -15
# L: -2
# Score: 44.0
# Length: 146
# Identity: 40/146 (27%)
# Gaps: 37/146 (25%)
```

### Fasta Format:

```
>seq1 1-131
```

```
>seq2 210-333
```

TAGTTATTAAACCTATCCAC - TCACTAGATACCCC - TTA - GCACA - -TAAAGGAGTATTATTGGGGTCAGCCATGGAGAAG - AATTTAGACAC - TAGGATAAGATAAGGAACACCCCAT - C - C - TAAAGA - A - A - T - C - A - C -  
 - - TCGGGCAGATACCCAAGACAATCAACAAGGGGAGA - GAATGACACAAGGAAGGACAGTCCAGTGCAACAAGCTAACGTGTGTGGC - - - AGGGGAAC - - - - - GTGACGGGGATTCTCC - AAAGATACACTACAAG

**3.3 Wnioski**