**Supplementary Material**

**Ribosomal DNA repeats in chicken and guinea fowl genomes**

Alexander Dyomin, Svetlana Galkina, Asya Davidian, Elena Gaginskaya**\***

**\*Corresponding author: E-mail: elena.gaginskaya@spbu.ru**

**Content:**

|  |  |
| --- | --- |
| **I. Supplementary Tables**…………………………………………………………...… | 2 |
| Table S1. Lengths and GC content of studied rDNA repeat elements of *Gallus gallus* (WAG137G4, AADN04001305.1) and *Numida meleagris* (JABXER010000123)……………………………………………………………… | 2 |
| **II. Supplementary Data**………………………………………………………….…… | 4 |
| Data S1.  One full rDNA repeat annotation of *N. meleagris* in JABXER010000123 contig…………………………………………………...…… | 4 |
| [Data S2.](https://docs.google.com/document/d/1CURXJ45lfTh02QJ-PxRiByUEYHWO2mEs3U_LEwNm44g/edit#heading=h.fbv9ash5b4ft) Consensus of *N. meleagris* IGS internal repeats………………………… | 12 |
| **III. Supplementary Figures**………………………………………………………...… | 14 |
| Figure S1. Identification of tandem repeat regions in the *N. meleagris* IGS (JABXER010000123) by dot-plot analysis……………………………...………… | 14 |
| Figure S2. VAL repeats in *N. meleagris...*………….……………………………… | 15 |
| Figure S3. Alignment of the 5S rRNA sequences found in the chicken and guinea fowl genomes……………………….……………………………………………… | 16 |

**I. Supplementary Tables**

**Table S1. Lengths and GC content of studied rDNA repeat elements of *Gallus gallus* (WAG137G4, AADN04001305.1) and *Numida meleagris* (JABXER010000123)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Gallus g. domesticus***  **(WAG137G4)**  **I** | ***Gallus g. domesticus***  **(WAG137G4)**  **II** | ***Gallus g. domesticus***  **(WAG137G4)**  **III** | ***Gallus gallus***  **(AADN04001305.1)**  **IV** | ***Numida meleagris***  **(JABXER010000123)**  **I** | ***Numida meleagris***  **(JABXER010000123)**  **II** | ***Numida meleagris***  **(JABXER010000123)**  **III** |
| **5’ETS** | 1,838 bp  74.8% GC | 1,841 bp  74.9% GC | 1,834 bp  74.8% GC | - | - | 1,792 bp  75.7% GC | 1,792 bp  75.8% GC |
| **18S** | 1,823 bp  54.4% GC | 1,823 bp  54.4% GC | 1,822 bp  54.3% GC | - | - | 1,824 bp  54.4% GC | 1,823 bp  54.4% GC |
| **ITS1** | 2,533 bp  82.2% GC | 2,528 bp  82.1% GC | 2,529 bp  82.2% GC | - | - | 2,837 bp  82.7% GC | 2,833 bp  82.6% GC |
| **5.8S** | 157 bp  57.3% GC | 157 bp  57.3% GC | 157 bp  57.3% GC | - | - | 157 bp  57.3% GC | 157 bp  57.3% GC |
| **ITS2** | 740 bp  82.2% GC | 734 bp  82.2% GC | 740 bp  82.2% GC | - | - | 761 bp  82.3% GC | 760 bp  82.2% GC |
| **28S** | 4,442 bp  68.0% GC | 4,412 bp  68.1% GC | 4,438 bp  68.0% GC | - | - | 4,422 bp  68.0% GC | 4,420 bp  68.0%GC |
| **3’ETS** | 338 bp  79.3% GC | 335 bp  79.4% GC | 335 bp  79.4% GC | - | - | 340 bp  83.2% GC | 340 bp  83.2% GC |
| **IGS** | 22,627 bp  68.1% GC | 15,169 bp  69.2% GC | 15,241 bp  68.1% GC | 14,002 bp  67.0% GC | 6,530 bp  68.1% GC | 6,562 bp  68.3% GC | 8,223 bp  70.4% GC |
| **poly-T** | 19 bp | 18 bp | 10 bp | 21 bp | 47  bp | 44  bp | 49 bp |
| **SV** | 137–158 bp | 139–158 bp | 139–157 bp | 128–169bp | 48–57bp | 48–56bp | 48–57bp |
| **AL** | 209–303bp | 209–290 bp | 209–290bp | 206–291bp | 156–236 bp | 151–236 bp | 161–231 bp |
| **Nme block** | - | - | - | - | 794 bp  67.1% GC | 838 bp  66.6% GC | 798 bp  67.4% GC |
| **UR** | 1,941 bp  64.0% GC | 1,937 bp  64.0% GC | 1,937 bp  64.0% GC | 1,959 bp  65.0% GC | 1,348 bp  69.4% GC | 1,354 bp  69.4% GC | 1,349 bp  69.2%GC |
| **CT(n)** | 33 bp | 33 bp | 33 bp | 33 bp | 32 bp | 36 bp | 36 bp |
| **IR** | 93 bp | 93 bp | 93 bp | 95 bp | 93 bp | 93 bp | 93 bp |
| **EL block** | 14,414 bp  71.4% GC | 9,668 bp  71.3% GC | 9,297 bp  71.3% GC | 7,057 bp  71.0% GC | - | - | - |
| **VAL block** | - | - | - | - | 574 bp  80.7% GC | 571 bp  80.4% GC | 571 bp  80.4%GC |
| **poly-A** | 7 bp | 7 bp | 7 bp | 7 bp | - | - | - |
| **UR** | - | - | - | - | 728 bp  77.5% GC | 754 bp  78.2% GC | 723 bp  77.6% GC |
| **VAL block (insertion)** | - | - | - | - | - | - | 975 bp  77.6% GC |
| **UR (insertion)** | - | - | - | - | - | - | 719 bp  77.2% GC |
| **VAL block** | 3,766 bp  76.4% GC | 2400 bp  75.6% GC | 2,400 bp  75.6% GC | 2,948 bp  75.9% GC | 1,671 bp  81.0% GC | 1,675 bp  81.1% GC | 1,674 bp  80.9% GC |
| **UR** | 190 bp  58.4% GC | 190 bp  58.4% GC | 191 bp  58.6% GC | 191 bp  58.6% | 160 bp  62.1% GC | 169 bp  62.1% GC | 169 bp  62.1% GC |
| **Full rDNA repeat** | 34,498 bp  69.1% GC | 27,000 bp  70.0% GC | 27,096 bp  69.4% GC | - | - | 18,695 bp  70.5% GC | 20,348 bp  71.2% GC |

**II. Supplementary Data**

**Data S1. One full rDNA repeat annotation of *N. meleagris* in JABXER010000123 contig**

LOCUS      JABXER010000123       18695 bp    DNA     linear   UNA 14-AUG-2022

DEFINITION  Numida meleagris breed Indigenous isolate YPT65100

            HiC\_scaffold\_123, whole genome shotgun sequence.

ACCESSION

VERSION

KEYWORDS WGS

SOURCE      Numida meleagris (helmeted guineafowl)

  ORGANISM  Numida meleagris

            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

            Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda;

            Coelurosauria; Aves; Neognathae; Galloanserae; Galliformes;

            Numididae; Numida.

FEATURES             Location/Qualifiers

     source          <1..>18695

                     /organism="Numida meleagris"

                     /mol\_type="genomic DNA"

                     /submitter\_seqid="HiC\_scaffold\_123"

                     /isolate="YPT65100"

                     /db\_xref="taxon:8996"

                     /sex="male"

                     /tissue\_type="blood"

                     /dev\_stage="adult"

                     /ecotype="Domestic-HGF, assemble"

                     /note="breed: Indigenous"

                     /label="breed: Indigenous source Numida meleagris"

     misc\_RNA        1..1792

                     /label="5'ETS"

     rRNA            1793..3616

                      /label="18S"

     misc\_feature    3617..6453

                     /label="ITS1"

     rRNA            6454..6610

                     /label="5.8S"

     misc\_feature    6611..7371

                     /label="ITS2"

     rRNA            7372..11793

                      /label="28S"

     misc\_RNA        11794..12133

                       /label="3'ETS"

     motif           12134..12177

                     /label="PolyT-region"

     misc\_feature    12134..18695

                     /label="IGS\_II"

     misc\_difference 12218..13323

                     /label="SV-AL"

     repeat\_region   12218..12265

                     /label="SV"

     repeat\_region   12266..12416

                     /label="AL"

     motif                12295..12301

      /label="PolyT-region"

     motif                12324..12330

      /label="PolyT-region"

     repeat\_region   12417..12470

                     /label="SV"

     repeat\_region   12471..12695

                     /label="AL"

     repeat\_region   12696..12751

                     /label="SV"

     repeat\_region   12752..12977

                     /label="AL"

     repeat\_region   12978..13031

                     /label="SV"

     repeat\_region   13032..13267

                     /label="AL"

     repeat\_region   13268..13323

                     /label="SV"

     motif           13335..13370

                     /label="CT-repeats"

     misc\_difference 13371..14172

                     /label="Nme"

     repeat\_region   13371..13478

                     /label="Nme"

     repeat\_region   13479..13669

                     /label="Nme"

     repeat\_region   13670..13851

                     /label="Nme"

     repeat\_region   13852..13983

                     /label="Nme"

     motif                14083..14099

      /label="PolyT-region"

     repeat\_region   14135..14172

                     /label="Nme"

     misc\_difference 14173..15526

                     /label="Unique region"

    motif                 14673..14687

      /label="PolyT-region"

     repeat\_region   15033..15125

                     /label="IR"

     misc\_difference 15527..16097

                     /label="Val Repeat Block"

     repeat\_region   15527..15618

                     /label="Val\_F"

     repeat\_region   15619..15702

                     /label="Val\_A"

     repeat\_region   15703..15787

                     /label="Val\_B"

     repeat\_region   15788..15872

                     /label="Val\_B"

     repeat\_region   15873..15956

                     /label="Val\_B"

     repeat\_region   15957..16041

                     /label="Val\_B"

     repeat\_region   16042..16097

                     /label="Val\_FC"

     misc\_difference 16098..16851

                     /label="Unique region"

     repeat\_region   16284..16375

                     /label="IR"

     misc\_difference 16852..18526

                     /label="Val Repeat Block"

     repeat\_region   16852..16926

                     /label="Val\_F"

     repeat\_region   16927..17010

                     /label="Val\_B"

     repeat\_region   17011..17094

                     /label="Val\_C"

     repeat\_region   17095..17180

                     /label="Val\_E"

     repeat\_region   17181..17265

                     /label="Val\_D"

     repeat\_region   17266..17349

                     /label="Val\_B"

     repeat\_region   17350..17434

                     /label="Val\_C"

     repeat\_region   17435..17520

                     /label="Val\_E"

     repeat\_region   17521..17605

                     /label="Val\_D"

     repeat\_region   17606..17689

                     /label="Val\_B"

     repeat\_region   17690..17773

                     /label="Val\_C"

     repeat\_region   17774..17859

                     /label="Val\_C"

     repeat\_region   17860..17945

                     /label="Val\_E"

     repeat\_region   17946..18029

                     /label="Val\_A"

     repeat\_region   18030..18113

                     /label="Val\_A"

     repeat\_region   18114..18199

                     /label="Val\_E"

     repeat\_region   18200..18283

                     /label="Val\_D"

     repeat\_region   18284..18370

                     /label="Val\_E"

     repeat\_region   18371..18456

                     /label="Val\_E"

     repeat\_region   18457..18526

                     /label="Val\_FC"

     misc\_difference 18527..18695

                     /label="Unique region"

     motif                  18666..18672

       /label="PolyT-region"

ORIGIN

        1 aggagagagt gcggactctg ctcctgaggc ggcggtgtag cgggcgctcg cgagcacgtt

       61 ggactggccg gcgagacacg aggcttgggc ggcggtggcg gcggtgtttt cgggcgtcgg

      121 ttcgttgcgt cgatccggtg gcctcggcga gttcctccgt ccccgtgccg ccgctcggtg

      181 gtctccggcc ggtccctcgg gcgaggggtg aaaaaaagcc ggcgaaggcg agagagaagc

      241 agggcgagga gaagctggct cccgcgaagc cgggccgtac gccgacgaga aggcgagagt

      301 gggaaaggga aagacagtcg aaagggaacc gtctcgcggg gccggtccgc cggcacccga

      361 acggacggac gaaggaacgc ggaagggccg ggggagtccc cggtgggggt cccgtcgcgc

      421 tcggtgaaga cgacgggagc cggagctcgt gggtcgtggc ggcggcgcct ctatctcctt

      481 tcgcacctca ttctcacccg cacgcgggag ccccgtctga tcctgtggcg ctcctcgggc

      541 gcgtcggaga ggcttcccgg cgggccggct ctcagggccg gtgttcaggc acgggcgagc

      601 acctctccaa gagcgttgcc cacgcacacc cacccgcact tgcgcgcgcg gggctttccg

      661 ccgcgccccc cgggaagggc tcgcgccttc cctcctcttc ttgccttttt tctgtcgatc

      721 gatgaggcca ctcgggtcgc gtcggagagg gcccccggcg ggccggcact tccgcgctcc

      781 cccgtcccgg ggcggcggcg gcggcggcgt ccggtgttca ggcacgggcg gcctcctctc

      841 cagctcgctt cccgtcgtcc gcgaggcgag gcgctcgccc gctggagcga gggcggcggc

      901 ggcggcggcg gcgtcttcgg ggcttgcagc ccggccgcgc tcgctcgtcc gcccgccccg

      961 cccgtccgtc cgccgcctcg gcaccgaagg gctcgcctga gcgaatccag ctgtgtgacg

     1021 gccgagcggc cccgcgagcc gcaggcgtac cgcgttcgtc gcgcgaggcg ccggcgccgg

     1081 cctcggttcg gggcccgtcg ggtgccggcc gcgagcagcg agccggcggg gtggcaaccg

     1141 aggaaaaccg gcaagggagc cgacgagaag ggaaaaggaa gcgcagacga gggggctgcg

     1201 cccgcccgag cgggcgagtc ccagagccgt gggttttcct gccgcgtccc cgcccgctgc

     1261 ggagcgtgcc gccccgggaa aagaaagaaa gggtccccgg tccgggggtc gcgcccgcct

     1321 cgaggtcgcg ttctcctcta gcacgtccgc cggtggcggc ggcggcggcg gcggggcttt

     1381 ccccgtccgc ccgcccgccc gtccgcctgt ccgtccgctt cgccgccggt cccggagggc

     1441 gggtcagccc cggccgtcgc ggcgccgcgt cgcgcgcgcg cgctcgcgtc cggctctcgc

     1501 gggagagccc ggagcgcgcc gcccgaaaag ctgcgcagcg gtccccgctc cttccccgcg

     1561 tgggggaggt cggcggggcc gccccggggg tcgggcgcgt ctcccgtcgt cggccgagcg

     1621 cgacggagag ccgccctccc cgccgagagg cgttcgcccc ggcggccgcc gtcgccgacc

     1681 ccgccgagaa agggaaagcc cgaagcgagc cgagccgagc gtgagagaga gagagagaga

     1741 gaggtcggcg gcgggccggg cccgtcgggt cgcgcgtgcc ccgtggcgcg gctacctggt

     1801 tgatcctgcc agtagcatat gcttgtctca aagattaagc catgcatgtc taagtacaca

     1861 cgggcggtac agtgaaactg cgaatggctc attaaatcag ttatggttcc tttggtcgct

     1921 cccctcccgt tacttggata actgtggtaa ttctagagct aatacatgcc gacgagcgcc

     1981 gacctccggg gacgcgtgca tttatcagac caaaaccaac ccgggctcgc ccggcggctt

     2041 tggtggactc tagataacct cgagccgatc gcacgccccc gtggcggcga cgacccattc

     2101 gaatgtctgc cctatcaact ttcgatggta ctgtctgtgc ctaccatggt gaccacgggt

     2161 aacggggaat cagggttcga ttccggagag ggagcctgag aaacggctac cacatccaag

     2221 gaaggcagca ggcgcgcaaa ttacccactc ccgacccggg gaggtagtga cgaaaaataa

     2281 caatacagga ctctttcgag gccctgtaat tggaatgagt ccactttaaa tcctttaacg

     2341 aggatccatt ggagggcaag tctggtgcca gcagccgcgg taattccagc tccaatagcg

     2401 tatattaaag ttgctgcagt taaaaagctc gtagttggat cttgggatcg agctggcggt

     2461 ccgccgcgag gcgagctacc gcctgtccca gcccctgtct ctcggcgccc cctcgatgct

     2521 cttaactgag tgtcccgcgg ggcccgaagc gtttactttg aaaaaattag agtgttcaaa

     2581 gcaggctggc cgccggaata ctccagctag gaataatgga ataggactcc ggttctattt

     2641 tgttggtttt cggaaacggg gccatgatta agagggacgg ccgggggcat tcgtattgtg

     2701 ccgctagagg tgaaattctt ggaccggcgc aagacgaact aaagcgaaag catttgccaa

     2761 gaatgttttc attaatcaag aacgaaagtc ggaggttcga agacgatcag ataccgtcgt

     2821 agttccgacc ataaacgatg ccgactcgcg atccggcggc gttattccca tgacccgccg

     2881 ggcagctccc gggaaaccca agtctttggg ttccgggggg agtatggttg caaagctgaa

     2941 acttaaagga attgacggaa gggcaccacc aggagtggag cctgcggctt aatttgactc

     3001 aacacgggaa acctcacccg gcccggacac ggacaggatt gacagattga gagctctttc

     3061 tcgattccgt gggtggtggt gcatggccgt tcttagttgg tggagcgatt tgtctggtta

     3121 attccgataa cgaacgagac tctggcatgc taactagtta cgcgaccccc gagcggtcgg

     3181 cgtccaactt cttagaggga caagtggcgt tcagccaccc gagattgagc aataacaggt

     3241 ctgtgatgcc cttagatgtc cggggctgca cgcgcgctac actgactggc tcagcttgtg

     3301 tctaccctac gccggcaggc gcgggtaacc cgttgaaccc cattcgtgat ggggatcggg

     3361 gattgcaatt attccccatg aacgaggaat tcccagtaag tgcgggtcat aagctcgcgt

     3421 tgattaagtc cctgcccttt gtacacaccg cccgtcgcta ctaccgattg gatggtttag

     3481 tgaggtcctc ggatcggccc cggcggggtc ggccacggcc ctgccggagc gtcgagaaga

     3541 cggtcgaact tgactatcta gaggaagtaa aagtcgtaac aaggtttccg taggtgaacc

     3601 tgcggaagga tcattaccgg ggccgaggcc gtggccgtgg ccgaggccgg gcgtccggcc

     3661 gcggccggca gtgccgtcgc gagcgccgca gcgctccgtc ttccgcgagc ccgctccgcg

     3721 cggagcgcgg ctcctccccg cgaaagcgag aaactcccaa gtcgctccgc gcgcccggcg

     3781 agagagagag aagcgagagc aagggcgcgg agcgccgccg ccgggccgcg tgcggggcgc

     3841 gccgagcccg cgggtgccgc gcgcgtcctc ctccctccgt cgtccgagcc ctcctcgtcc

     3901 gtccgtctcg tccgcccgcc ccgccccgtc ccgtcccgcc gctcgtcgcc tcacggaggc

     3961 gagcgagcga gcgagggcga gggccgcggg gtctggtcgg cgggtcggcg ggcggtggcg

     4021 cgtggccggg gggtgtgggg gcgcgctccg ggcccgcggt gcggcgcccc gccgccccgg

     4081 gcgccggtcc gccgccggtc cgcccgacgc cgggtccggt ccgtccggtg tcgtcctcgc

     4141 tgccgtgccg tgccgcgccg tgccgcgcgc cggccggctg ccgcgcgctc gcgcgcgttc

     4201 ctccccgagc ctcgcctcgc cttggtcgcc gctctccggg gcgccggtcc cgcgtccgtc

     4261 ccgctccgac cgcggggtcg gggctgggcg gcggcggcgg gtgggggggg ggggaaggac

     4321 ggctctcccg tcccctcgct cggccggaag cccgccgccg gccctcgccg ccgccgccgc

     4381 cggcaccccg agtccgctcg agggaaagcc gcgcggacgg cggcgcgggg gcaggccgcg

     4441 ggcggcgggc tcgagcgtgc agggggcgag aagccggccg cttcccccgg cccccccccc

     4501 agcccgctct ttccccggcc cccgcgggga ggacggggtc ggtccgcggg cggccgcgga

     4561 gccgcggtcg accgaccgag ggcgggcgcc gagggggctc cgccgcgttc gcgcggcagc

     4621 gggcgggcgg ccgagcgagc cggcagcgag gcctccgggc ggcttgggtc ccgcgcgggc

     4681 ggtccgagcc gcgtttctcc tcccgggcgc agctgccgcc gggcgccggg ttaccgaggg

     4741 aaaccccggg ccccgggatg gaggacgagg tggtggcggc gggcgacggg cgcgcccccg

     4801 cgggcggacg ctcccccgag gggcgccggg gccggctggc gggtgccggg tctcccctcg

     4861 gcgccccgcc ccgccccgct ccgctccgct cggccggccg cggccggcgg gcgggcgggc

     4921 gggcggggga ggcacccccg cggggccttc gggtcgtttc cctcacccca gggccaggta

     4981 cctagcgtcc gcgcctctgc gcgcacgtcc ggtggttgcg ggggctgggg accggcaccg

     5041 agcgggccgc gttcccccca ccccgacccc gacgtcggcc gcagggccgg gcggaggttt

     5101 aaagactcgg gcggcccgcg gcgcgtcccg ccaggggcgg ggccgggggc ggtcggcggg

     5161 ccggcgggcc gccgggaagg gctcggttcg gtgggggggg gggcgcggcg gcgcgcccct

     5221 cccctccctc tcccctgccg tctcccttcc ccggcggtcc cgccggcggc tgccgccgcc

     5281 gccgcgctct cgctcctcgg cgggcgggcc cgggccggag aggggttcta ccctgtcccc

     5341 cctctccgcg gcctcggtct cgggcagaga gggttgacgc gcgcgcgcgc gcgcgcgctc

     5401 gcgcgcccgc gccggctccg gccgatccga tccgagccgc tggcgttcgg agcgggccga

     5461 ggcgcgggtc tcggccccgc cgctcctcgc ggcggcggct gccttcctgg cgcggcctcc

     5521 cctccccgcc tttcggccgt cttccgcgcg cgcggccgag gcgcctcgcg ccgcctccgc

     5581 cgggcgtcgc cgccgctcgc cacccggccg agccccgccg cacgccgcgt cctctccttc

     5641 gtctccgccg ccgccgccgc ctcccaccgc gcttgtcgcc ctcggccccg ccggccgcgc

     5701 cgggcgagcg ggcgggcggt ccggcgtggg gcgtcccgca gccggtctcc gcgcggaggc

     5761 gcggggagcg ggcgccgctc cgggatccgt ccccgtcccg cccgccgccg ggcgcgcgtc

     5821 cgccgcgggc gcgccgccag ggcgagcgag aggaggaggc ggcggagggc gcggggcggg

     5881 agaggagggc gaaaggcgag gctcgcgtcg tggggtcggc gcgcgcggcg cccggagggg

     5941 gggcgaggcc gaggcgtctc ggtcggtcgc ggtcggtcgg ccggccggcc ggcggcggag

     6001 gtccgccgcg gagggagggc gccgtgccgt gccgtgccgt gccggtccgg cgcgggccgg

     6061 ggcggcgcgt cgtgccgcgc gcgtctcctc ctcctgttct cgggccgagc cgagccgggc

     6121 gcctggtccg gccccgaagc gagacagggc cgtttcccca gggtgggagc gagggctccc

     6181 cggcccttct cgttcgggtg cgctgtggcg gccgcccacg ggcggctccc ttccccactc

     6241 gttgtcgtgt gctggtgcgg tcagcggagg cgacgctcct ccgccccgcg gtcggtcggt

     6301 cggtcggtcg tcccggcgtc ggggcgggcc gggccgggcc gggccgcggg gcgtcactcg

     6361 ggaggagcgc cttcgggaaa gaaaggcgag agagagagag agagagaaac gagagggcac

     6421 ccccgcgcgc gcgggccggt gcggcaagtc agacaactct tagcggtgga tcactcggct

     6481 cgtgcgtcga tgaagaacgc agctagctgc gagaattaat gtgaattgca ggacacattg

     6541 atcatcgaca cttcgaacgc acttgcggcc ccgggttcct cccggggcta cgcctgcctg

     6601 agcgtcgctt gacggtcaat cgccgagggc cgccgtccgc cgccgccgcc tccggggcgg

     6661 cggcgaaggc ggcggccgcg cggcgcggct ggggcgcctc gcaggcccgc gcgtcccgcc

     6721 ggtgggcggg gtcgggggtc cgtccttccg tccgcccggt cggtcggtcg gtcgtccggt

     6781 acccccgccg ccccgcccgg cggcggtggc gcgccgcggg ccttcgtccc cctaagtcga

     6841 gacccaggtc ggggagctcg cccggcgagc gagctccccg cgctcccgga gcgcccgctt

     6901 gggccgagct cgtccccacg gggcggccgg gctttccggg tcggtcgcgc ggtcgggggc

     6961 cggtcggtcg gttgtccggc cggccggccc tcggcccgga ggagcgttcc gcgccctccc

     7021 cggccccgcg cgcggctgcc tgcgggtcgc gttaccggcg gcggtagcgc gccgtgctgc

     7081 cgcgcgcgcg cggcgctgcg gggggggcga ggcgcgtccg ccctccggtc gtgcgtccgt

     7141 ccgtccgtct cgcccggcga gtccccgcgg gtccgtctct ggccgccggc ggcggcggtg

     7201 gcggcggcgg cggcggcggc ggcggtcgtt ccttcgtcgg tccgtccgtc cgtccgtcgg

     7261 ctccccgccc gcccgcccgc cgtcggttcc gtccgggcgt cctccgcgtc gggagtctgc

     7321 ttcgcctccg agcggggaac cgcgaggttt cccggcggcc ggctttccga tcgcgacctc

     7381 aggtcagacg tggcgacccg ctgaatttaa gcatattagt cagcggagga aaagaaacta

     7441 acgaggattc cctcagtaac ggcgagtgaa gagggaagag cccagcgccg aatccccgcc

     7501 ccgcggtggg gcgcgggaag tgtggcgtac ggaagccccc atccccggcg ccgctctcgg

     7561 ggggcccaag tccttctgat cgaggcccag cccgcggacg gtgtgaggcc ggtagcggcc

     7621 ccccggcgcg ccgggcccgg ggcttctcgg agtcgggttg cttgggaatg cagcccaaag

     7681 cgggtggtaa actccatcta aggctaaata ccggcacgag accgatagcc aacaagtacc

     7741 gtaagggaaa gttgaaaaga actttgaaga gagagttcaa gagggcgtga aaccgttaag

     7801 aggtaaacgg gtggggtccg cgcagtcggc ccggaggatt caacccggcg ggccaaggtc

     7861 ggccggcgcg ggcgccgtcg gatccccgcc tccgcctccc ctccgtcctt ccccttcgcc

     7921 ggggcggggc gggcccaggg ggggcgggcg ggccggggac cgccgcccgg ccggcgtccg

     7981 gcccccgtcg ggcgcatttc ctccgcggcg gtgcgccgcg accggctccg ggacggccgg

     8041 gaagggctgc cggcgggcag gtggcccggc gccgcgcgag cggcggcggg tgttagagcc

     8101 gccgggcccc gacggtcgcc gaatcccggg gccgagggag aggaccgccg ccgcgccctc

     8161 ccccggaggg gggcggcccc ccggagggtc cctcgccgga ccggcgtcgg gccggccgcg

     8221 cctcgccgcc gccgccgccg ccgccgccgc tctctcgttc cccgcctcgg gtccgcccgg

     8281 ggcgcggggg cggggggggt ggccggcgcg cgagcgagcg gcgcggcgcg cgcctccagc

     8341 ccggcgcggg cgaggcgggg gccgccgggg gggacgaacc ttcccccttc cctggctcgg

     8401 gcccgctcct cgctcctggc ggggagcggc gcccgttccg ctcgggggac gggcccgccg

     8461 gcccccggcg ccgctgtttc cgaccggggc ggactgcgct cagtgcgccc cgaccgcgcg

     8521 gcgccgccgg gccgggctcg ggccacgccg aggcgctcgg ggtccgcggc gacgtcggct

     8581 acccacccga cccgtcttga aacacggacc aaggagtcta gcacgcgcgc gagtcggcgg

     8641 ctcgcgcgaa agcccgcggc gcaatgaagg tgagggccgg cgcgcgccgg ctgaggtggg

     8701 atcccggggc ggcaggccgg aaggccccgg gcgcaccacc ggcccgtctc gcccgcctcg

     8761 ccggggaggt ggagcatgag cgcgcgtgct aggacccgaa agatggtgaa ctatgcctgg

     8821 gcagggcgaa gccagaggaa actctggtgg aggtccgtag cggtcctgac gtgcaaatcg

     8881 gtcgtccgac ccgggtatag gggcgaaaga ctaatcgaac catctagtag ctggttccct

     8941 ccgaagtttc cctcaggata gctggcgctc tcggtggcgg ggcagtttta cccggtaaag

     9001 cgaatgatta gaggtcttgg ggccgaaacg atctcaacct attctcaaac tttcaatggg

     9061 taagacgccc ggctcgctgg cgtggagccg ggccgtggaa tgcgagcgct cagtgggcca

     9121 cttttggtaa gcagaactgg cgctgcggga tgaaccgaac gccgggttaa ggcgcccgat

     9181 gccgacgctc atcagagccc agaaaaggtg ttggttgatc tagacagcag gacggtggcc

     9241 atggaagtcg gaacccgcta aggagtgtgt aacaactcac ctgccgaatc aactagccct

     9301 gaaaatggat ggcgctggag cgtcgggccc atacccggcc gtcgccggcg gtgcggagcc

     9361 gcgggggcta cgccgcgacg agtaggaggg ccgctgcggt gcgcctggaa gcctggggcg

     9421 cgggcccggg tggagccgcc gcaggtgcag atcttggtgg tagtagcaac tattcaaacg

     9481 agagctttga aggccgaagt ggagcagggt tccatgtgaa cagcagttga acatgggtca

     9541 gtcggtccta agcgataggc gagcgccgtt ccgaagggac gggcgatggc ctccgttgcc

     9601 ctcggccgat cgaaagggag tcgggttcag atccccgaat ccggagcggc ggagacgggc

     9661 gccgcgaggc gcccagtgcg gtaacgcaag cgatcccgga gaagccggcg ggagccccgg

     9721 ggagagttct cttttctttg tgaagggccg ggcgccctgg aacgggttcg ccccgagaga

     9781 ggggcccgcg ccttggaaag cgtcgcggtt ccggcggcgt ccggtgagct ctcgctggcc

     9841 cgtgaaaatc cgggggagag ggtgtaaatc tcgcgccggg ccgtacccat atccgcagca

     9901 ggtctccaag gtgaacagcc tctggcatgt tggaccaatg taggtaaggg aagtcggcaa

     9961 gccggatccg taacttcggg ataaggattg gctctaaggg ctgggtcggt cgggctgggg

    10021 cgcgaagcgg ggctgggcgc gcgccgcggc tggacgaggc gccgcccgcc cccgcctccc

    10081 ttcccccgcc cccccgctcg ccggggggcg ccgggggggg ggtgccggcg ggcggcgcgg

    10141 cggcggcgac tctggacgcg cgccgggccc ttcccgtgga tcgccccagc tgcggcgggc

    10201 gccgcccgcc cccctccttg cccctccgcc cctccgcccc cggcgccctc ccgcccgtcc

    10261 cccccgctcc cctccctccg cccgagccgc ccgctccgcc ggcgcggcga ggcgcggggg

    10321 ggccccgggg aggcgcgggg ggctccgggc ggggggggcc ccgccggcgg cgccgggggg

    10381 gcggtcccgg gcgggggggg tctccgggcc ggcgccccgc ctcggccggc gcctagcagc

    10441 cggcttagaa ctggtgcgga ccaggggaat ccgactgttt aattaaaaca aagcatcgcg

    10501 aaggcccgcg gcgggtgttg acgcgatgtg atttctgccc agtgctctga atgtcaaagt

    10561 gaagaaattc aatgaagcgc gggtaaacgg cgggagtaac tatgactctc ttaaggtagc

    10621 caaatgcctc gtcatctaat tagtgacgcg catgaatgga tgaacgagat tcccactgtc

    10681 cctacctact ctccagcgaa accacagcca agggaacggg cttggcggaa tcagcgggga

    10741 aagaagaccc tgttgagctt gactctagtc tggcgctgtg aagagacatg agaggtgtag

    10801 aataagtggg aggccccgcg gtcgcgcgac ccgcgccgcg gcccggccgc cggtgaaata

    10861 ccactactct gatcgttttt tcacttaccc ggtgaggcgg gggggcgagc cccgaggggc

    10921 tctcgcttct ggcgccaagc gggcggctca agcgcgccgg ccgcgacccg ctccggggac

    10981 agcgtcaggt ggggagtttg actggggcgg tacacctgtc aaagcgtaac gcaggtgtcc

    11041 taaggcgagc tcagggaggc cagaaacctc ccgtggagca gaagggcaaa agctcgcttg

    11101 atcttgattt tcagtacgaa tacagaccgt gaaagcgggg cctcacgatc cttctgactt

    11161 tttgggtttt aagcaggagg tgtcagaaaa gttaccacac agggataact ggcttgtggc

    11221 ggccaagcgt tcatagcgac gtcgcttttt gatccttcga tgtcggctct tcctatcatt

    11281 gtgaagcaga attcaccaag cgttggattg ttcacccact aatagggaac gtgagctggg

    11341 tttagaccgt cgtgagacag gttagtttta ccctactgat gatgtgttgt tgcgctagta

    11401 atcctgctca gtacgagagg aaccgcaggt tcagacattt ggtgtatgtg cttggctgag

    11461 gagccactgg agcgaggcta ccatctgtgg gattatgact gaacgcctct aagtcagaat

    11521 cccccctaaa cgtagcgata ccgcagcgcc gaggcgcctc ggtgggctcg cgatagccgg

    11581 ccgcgtggcc ggtgcggagc gccgctcgtg gtcgggaccg gagcgcggcc ggatgcggcg

    11641 ccgcctctcc cccgccgcgt accgcacgtt cgtggggcac ccggtgctaa atcattcgta

    11701 gacgacctga ttctgggtcg gggtttcgta cgtagcagag cagctccctc gctgcgatct

    11761 attgagagtc agccctcgac acaagctttt gtcggccggg ggaggcgagg ctcgccgcgg

    11821 cgcggcgcgg ctccggcgcg ggccggggcc ggggccgggg ccgggcgggc ggcgctcgtc

    11881 tcttcccctt ggtcttcctc cgggtcctcc tctgcccgct cgcctgcgcc gccggcgagg

    11941 cgaggggagg ggagggggaa ggggacgcgg ttccgcgcgc gcgggcgcgc gccggtcccc

    12001 ttccttccgt ttcctcccgc cggagccggg gggttgacct ggcgggccgg ggaggggcgg

    12061 gcggtcgtcc cgctgggcgg cggggcagcc gcccgaccca cccgctcggt cgggggttga

    12121 cctggcggcc ggcttttttt tttttttttt tttttttttt tttttttttt ttcttttccc

    12181 ccttgggagc cgtcggctcg ccgatgggcc tcccgcccgg gaagaccacc cctgacgaga

    12241 cggagcgggt agacctggcg ggcggttatt tattcctccc cccccccccc ccgctttttt

    12301 taaaaatatt ttaattttta ttcttttttt attaattcat tcgtccgtcc gcccgttcca

    12361 gttcttggcg agccgagccg agccgagccg agccgagccg agccgagccg agccgagggt

    12421 cagacgaggc ggccgtgacg acgatggagc gggtagacct ggcgggcggg tttttttttt

    12481 ttttttcttt ttttaatttt ttttcttttt ttaatttttt tattcgattt tttatttttt

    12541 tattttttaa tttttttaaa ttttattttt tactcaattt ttcatttttt ttaatttttt

    12601 taaatttttt aattcgattt tttttgttgt tttaaatatt tatttattta tttatttatt

    12661 caagcgtcca tccgtccgtt cctgttcttg gcgaggggtc agacgaggcg gccgtgacga

    12721 cgatgggagc gggtagacct ggcgggcggg gttttttttt tttgtttttt tttttttgtt

    12781 tgttttttgt agtttttttt atttttttaa attttgtttt ttaactcgat tttattttta

    12841 tatattaatt ttttaattcg attttttatt tttttaattt ttttaaattt tttaaattat

    12901 tttatttata tatttattca agcgtccgtc cgcccgttcc agttcttggc gagccgagcc

    12961 gagccgagcc gagccgaggt tcagacgagg cggccgtgac gacgatggag cgggtagacc

    13021 tggcggccgg cttttttttt attatttttt ttaatttttt ttttaatttt tcttttttct

    13081 ttaatttttt aaaatttttt aattcgaatt atttattcat tttatttata tatatttttt

    13141 taatttttta aattatttta tttatttatt tattcaaccg tccgtccgtc cgtccgtccc

    13201 agttcttggc gagccgagcc gagccgagcc gagccgagcc gagccgagcc gagccgagcc

    13261 gagccgaggg tcagacgagg tggccgtgcc gatggagcgg gagggtcaga cctggcggcc

    13321 ggtgcctgag cgcgctctct ctctctctct ctctctctct ctctctctct ccccccctcc

    13381 caacccgccc cccccccccc cccccccccc cccgaccccc ccgattgttc cggttttcga

    13441 catctctgtt ttaaacacgt gggttttttt ttttccttcc cccccccccc cccccccccc

    13501 ccgcctttca gggacccctc cccccaccca ccccggcggg gcaggcccga tggctcggcc

    13561 ggaccgtcgg catgacccga cggacagagg ccgcccgccc cccaaccgac ccctccgatt

    13621 gttccggttt tcgacatctc tattttaaac acgtggggtt ttttttcctc cccccccccc

    13681 ccccgccttt cagggacccc tccccccacc caccccggcg gggcaggccc gatggctcgg

    13741 ccggaccgtc ggcatgaccc gacggacaga ggccgcccgc cccccaaccg acccctccga

    13801 ttgttccggt tttcgacatc tctattttaa acacgtgggg ttttttttcc tccccccccc

    13861 cccccccccc gcctttcagg gacccctccc cccacccacc ccggcggggc aggcccgatg

    13921 gctcggccgg accgtcggcg cgacccgacg gacagaagcc gcttgccccc caacccccca

    13981 accgcttcgc cccgtcggcc cggttcatgt tcctttgtct tttccccggg gagggggggg

    14041 gagcggaggg gaaggggggg cgagggggag ttctttattt agtttttttt ttttttttta

    14101 tgaataaaac ccccgctttc atccgacggg ggatagacct gatggctcgg ccggaccgtc

    14161 ggcgtgactc gaggtctggc gcccctaccc ctctcttcac cccccccccc cccccccccc

    14221 cccccccgga tcgcgcccgc ccgccctgat cgggccccgc ttttctatac atacattttt

    14281 aaagtttctt ctgcttcttc ttctctctct ctctctctct ctctctctct ctctctctct

    14341 ctctctctct ctcccccccc ccctctctct ctctctcgca gacctcgtgc ctcggccgga

    14401 cggcggtggc gagactcggc ggcccccgcc cccgtggagc caccgcggcc cgccgtcccg

    14461 ccgctttgtt tcgtttctcc ggacggctgg ctttgccttg cccttcttcc gaccgagaga

    14521 gcgaccgacc gtcctccgcg aggcgagctg ccctgccgcc ggccgggaga ggagctcggc

    14581 cagaggacgg ctcgaacgaa cccgcgaccc agacccccgc cgccccgtct cacgccggcc

    14641 tacctctgct gcctgcctcc cgtcttcccg cctttttttt tttttttccc gcgccgaccc

    14701 ggcccgagcg agcgagcgag cgagcgagcg agcgagcccg tccgtccgtc cgtccgtccg

    14761 tccgctggct gaatccgacg cgctttcccg cccacccggc ccctgctggg cggggtcccc

    14821 gcccgcgcct gcgcccttcc cccggaggtc gggggcgcgc ccaagcccgc tcccaagccc

    14881 gtgcccgcgc cgtttcttcc agaacggccg aagcgtagaa tctgcctagg gcgggaaaca

    14941 cctccaacgg tcgtccggcg accacggtcc acccctcacc ggcgcaccgc accgccccct

    15001 ccaccctccc ccccacctcc ccccgccccc gagcccccgc gagccgaacg cggattagct

    15061 ttccaatatt cgacggaaaa ccacgggaag agttacctag ccaccgagcg aaaaagtaaa

    15121 cggaaataaa taatctttta aaaagtacga aaacgagaac aaacagtcaa agaagcgagc

    15181 gagcgagcga gcgagcgagc gagcaaacaa accgacgaga agagtaaaac ccaagagcaa

    15241 cagggagact gaagggacgg tctcgcgccg cggccggcca gccggccgcc gaaagcgtcc

    15301 tcttggggcg agcaggggcc gccgccgcct cgacccggcc ccgagccggc ggccgcgaga

    15361 tctgcacgcc gctgagcgtc ggcgcccggc tctgccggtg cccagagcga gggagggagc

    15421 gagcagcggc ggcgttgcgg cggcggcgga ggcggaggcg ttgcggtcgg acggccggcg

    15481 gggccagagc ttcccgtgcc ccgagcctcg gcgagtcggc tcggctcggc tcggttcggc

    15541 ccggccccga gcggctgcgc gcgggtccgc ccggccgcga acgccggctc cggccgggct

    15601 cgaccgctcc gcgcagcccg gctaggggtc gctgccgcgg cggcttgccg agcccggcca

    15661 cgtctacccg gctccggccg ggctcgcccg ccccgcgccg cccgcctagg ggtcgctgcc

    15721 ggggcggctt gcatcgttcg gccacgtcta cccggctccg gccgggctcg aaccgccccg

    15781 cgctgcccgg ctaggggtcg ctgccgcggc ggcttgcatc gttcggccac gtctacccgg

    15841 ctccggccgg gctcgaaccg ccccgcgccg cccggctagg ggtcgctgcc gcggcggctt

    15901 gcatcgttcg gccacgtcta cccggctccg gccgggctcg cccgccccgc gcagcccggc

    15961 taggggtcgc tgccggggcg gcttgcatcg ttcggccacg tctacccggc tccggccggg

    16021 ctcgaaccgc cccgcgctgc ccggctaggg gtcgctgccg ggcggcggct tgcctcgctc

    16081 ggccacgtct acccagcggg gggggggggg gggggggggg gggggggggg gggggggggg

    16141 gggggggggg ggcggggggg gcggacggca cggcgcgctc gctcgaccgc tcgccttgga

    16201 aaacgcagca gcagcgcgcg ccgcccggcg agcaccccct acctccctcc cacccccccc

    16261 tcaccgcccc gctccgcccc tgtgcccccg cgagccgaac gcggattagc tttccaatac

    16321 tcggcggaaa actccgaaag agatgcccgg ccaccgagcg aacaagtcat cggaagtgca

    16381 attcaaagac ctttaaaaaa gaacgaaaac gagaacaaac agtcaagcac tcggctcggc

    16441 acggcgcggc acggctcggc tcggctcggc ccggccccgc cccgagcggc tgcgcccggg

    16501 tccgcccggc cgcgaacgcc ggcggccgcc ccgacccgtg ccggcgttgc ggcggcgggg

    16561 gcggcctcga cccggccccc gagcggcggc agcggcggcc gcccgggtgg accacggtcc

    16621 gagccgcggc ggccacccgg gtgcgcgctc ggccgagcgg cggcggtcat gtcggcgcgc

    16681 cgccgagccg cggcggccgg cgcgagcgag atcctccgga ctcgcgcgcc gcgaccggcg

    16741 gcttctgcat gccgctgccg gagcgcccag ggaacgccgg ccggctctcc ctgcctggct

    16801 ccggccgaca cgcgagcccc agccggcggc ggcggaagga cgccgccgga gcggctcggg

    16861 atcgccggcc ccggagctct ccggcatcgg ctcgtcccgt cccgccccgc cccgccccgc

    16921 gcagcccggc taggggtcgc tgccgcggcg gcttgcatcg ttcggccacg tctacccggc

    16981 tccggccggg ctcgaccgcc ccgcgccgcc cggctagggg tcgctgccgc ggcggctagc

    17041 ctcgctcggc cacgtctacc cggctccggc cgggctcgcc cgccccgcgc cgcccgccta

    17101 ggggtcgctg ccgggcggcg gcttgccgag cccggccacg tctacccggc tccggccggg

    17161 ctcgcccgcc ccgcgccgcc cggctagggg tcgctgccgc ggcggcttgc cgatcccggc

    17221 cacgtctacc cggctccggc cgggctcgaa ccgccccgcg ccgcccggct aggggtcgct

    17281 gccgcggcgg cttgcatcgt tcggccacgt ctacccgtct ccggccgggc tcgaccgctc

    17341 cgcgcagccc ggctaggggt cgctgccgcg gcggctagcc tcgctcggcc acgtctaccc

    17401 ggctccggcc gggctcgaac cgccccgcgc cgcccggcta ggggtcgctg ccgggcggcg

    17461 gcttgccgag cccggccacg tctacccgtc tccggccggg ctcgcccgcc ccgcgccgcc

    17521 cggctagggg tcgctgccgc ggcggcttgc cgatcccggc cacgtctacc cggctccggc

    17581 cgggctcgaa ccgccccgcg ccgcccgcct aggggtcgct gccgcggcgg cttgcatcgt

    17641 tcggccacgt ctacccggct ccggccgggc tcgaccgctc cgcgcagcac ggctaggggt

    17701 cgctgccgcg gcggctagcc tcgctcggcc acgtctaccc gtctccggcc gggctcgccc

    17761 gccccgcgct gcccggctag ggatcgccgc cgcggcggct tgcctcgctc ggccacgtct

    17821 acccgtctcc ggccgggctc gaccgcgccc cgcgccgccc gcctaggggt cgctgccggg

    17881 cggcggcttg ccgaccccgg ccacgtctac ccggctccgg ccgggctcga ccgccccgcg

    17941 cagcccgcct aggggtcgct gccgcggcgg cttgccgagc ccggccacgt ctacccgtct

    18001 ccggccgggc tcgaccgccc cgcgcagccc ggctaggggt cgctgccgcg gcgtctagcc

    18061 gagcccggcc acgtctaccc gtctccggcc gggctcgccc gccccgcgca gcccggctag

    18121 gggtcgctgc cgggcggcgg cttgccgagc ccggccacgt ctacccgtct ccggccgggc

    18181 tcgaccgctc cgcgcagccc ggctaggggt cgctgccgcg gcggcttgcc gatcccggcc

    18241 acgtctaccc ggctccggcc gggctcgccc gccccgcgca gcccggctag gggtcgctgc

    18301 cgggcggcgg cttgccgagc ccggccacgt ctacccggct ccggccgggc tcgaaccgcc

    18361 ccgcgcagcc cggctagggg tcgctgcctg gcggcggctt gccgagcccg gccacgtcta

    18421 cccggctccg gccgggctcg accgccccgc gctgcccggc taggggtcgc tgccgggcgg

    18481 cttgcctcgc tcggccacgt caaccgggct ccgtccgccc ttcgccgtgt tggcgtgggg

    18541 ggggcgctgt cgcttcggag ctgccgtagg ggtcgctgcc attttttgcc tcccgttacg

    18601 tcaaggtaaa ccccggctgc tgtccgagca gctgccggta gtcggcgccg atggctttgg

    18661 aggccttttt ttgggtgctt tccctgtccg tctat

//

**Data S2. Consensus of *N. meleagris* IGS internal repeats**

**REPEAT REGION SV-AL**

**SV**

Consensus (threshold – 75%):

GGGTCAGACGAGGNGGCCGTGACGANGAHGNGAGCGGGTAGACCTGGCGGSCGGBG

L = 48 - 57 bp

GC% ≈ 74

**AL**

Consensus (threshold – 85%):

 TTTTTTTTTTTTTTTTSTTTTTTTWWTTTTKTTTSTTTTTTKWAKTTTTTTWTWNNNTTTTTTWWWWTTTTDTTTTTWAWYTYKNTTTWWWYTTTAWWTWTTWNNNAWTTTTTMATTYKWWTTATTTATTNNTNNNNNTHTTTANNNNNATTTTTTTANNNATTTTTTAAATTWTTTTATTTATTTATTTATTCAANNNCGTCCGTCCGYCCGTTCCAGTTCTTGGCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGA

L = 151 - 236 bp

GC% ≈ 28

**REPEAT REGION Nme**

**Nme**

Consensus (threshold – 85%)

CCCCCCCCCCCCCCCCCCCCCCCCCCGCCTTTCAGGGACCCCTCCCCCCACCCACCCCGGCGGGGCAGGCCCGATGGCTCGGCCGGACCGTCGGCRYGACCCGACGGACAGAVGCCGCCCGCCCCCCAACCGACCCCYCCGATTGTTCCGGTTTTCGACATCTCTRTTTTAAACACGTGGGKTTTTTTTTCCTT

L = 38 – 193 bp

GC% ≈ 70

**INTERSPERSED REPEAT (IR)**

Consensus (threshold – 85%)

GCCCCCGCGAGCCGAACGCGGATTAGCTTTCCAATAYTCGRCGGAAAACNWCSGRAAGAGWTRCCYRGCCACCGAGCGAAMAAGTMAWCGGAA

L = 92 – 93 bp

GC% ≈ 57.2

**REPEAT REGION VAL**

Consensus (threshold – 85%):

CGSCTAGGGGTCGCTGCCNNGCGGCGGCTWGCMKMKYYCGGCCACGTCTACCCGKCTCCGGCCGGGCTCNRMCCGCYCCGCGCMGCC

L = 54 - 92 bp

GC% ≈ 80.8

**Val\_F**

Consensus (threshold – 85%):

CGGCTCGNGWTCGNNCCGGCCCCNGAGCNKCTSCGNGCRNNTCSGCYCGKCNNNNNNNNCCGNNTCCSGCCSSGCNCNNCCGCYCCGCGCAGCC

L = 75 – 92 bp

GC% ≈ 94.9

**Val\_A**

Consensus (threshold – 85%):

CGSCTAGGGGTCGCTGCCGCGGCGKCTWGCCGAGCCCGGCCACGTCTACCCGKCTCCGGCCGGGCTCGMCCGCCCCGCGCMGCC

L = 87 bp

GC% ≈ 82.3

**Val\_B**

Consensus (threshold – 85%):

CGSCTAGGGGTCGCTGCCGSGGCGGCTTGCATCGTTCGGCCACGTCTACCCGKCTCCGGCCGGGCTCGANCCGCYCCGCGCHGCM

L = 84-85 bp

GC% ≈ 78.8

**Val\_C**

Consensus (threshold – 85%):

CGGCTAGGGRTCGCYGCCGCGGCGGCTAGCCTCGCTCGGCCACGTCTACCCGKCTCCGGCCGGGCTCGNNCNCGCCCCGCGCYGCC

L = 83-86 bp

GC% ≈ 81.6

**Val\_D**

Consensus (threshold – 85%):

CGGCTAGGGGTCGCTGCCGCGGCGGCTTGCCGATCCCGGCCACGTCTACCCGGCTCCGGCCGGGCTCGNMCCGCCCCGCGCMGCC

L = 84-86 bp

GC% ≈ 81.4

**Val\_E**

Consensus (threshold – 85%):

CGSCTAGGGGTCGCTGCCGGGCGGCGGCTTGCCGASCCCGGCCACGTCTACCCGKCTCCGGCCGGGCTCGMCCGCCCCGCGCMGCC

L = 86-87 bp

GC% ≈ 82.6

**Val\_FC**

Consensus (threshold – 85%):

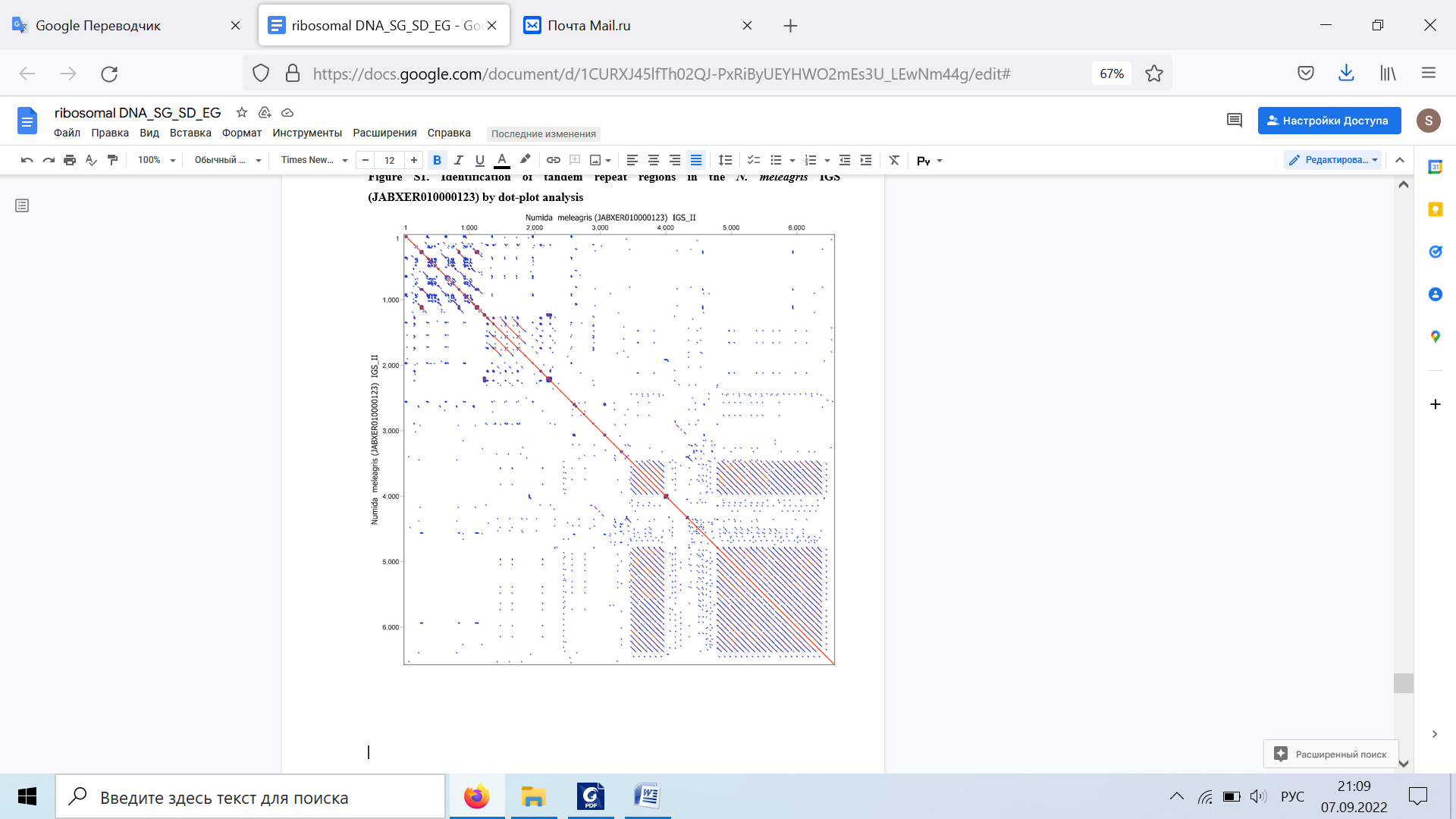
CGGCTAGGGGTCGCTGCCGNNNGGCGGCTTGCCTCGCTCGGCCACGTCWACCSRGCTCCGTCCGCCCTTCGCC

L = 54 – 70 bp

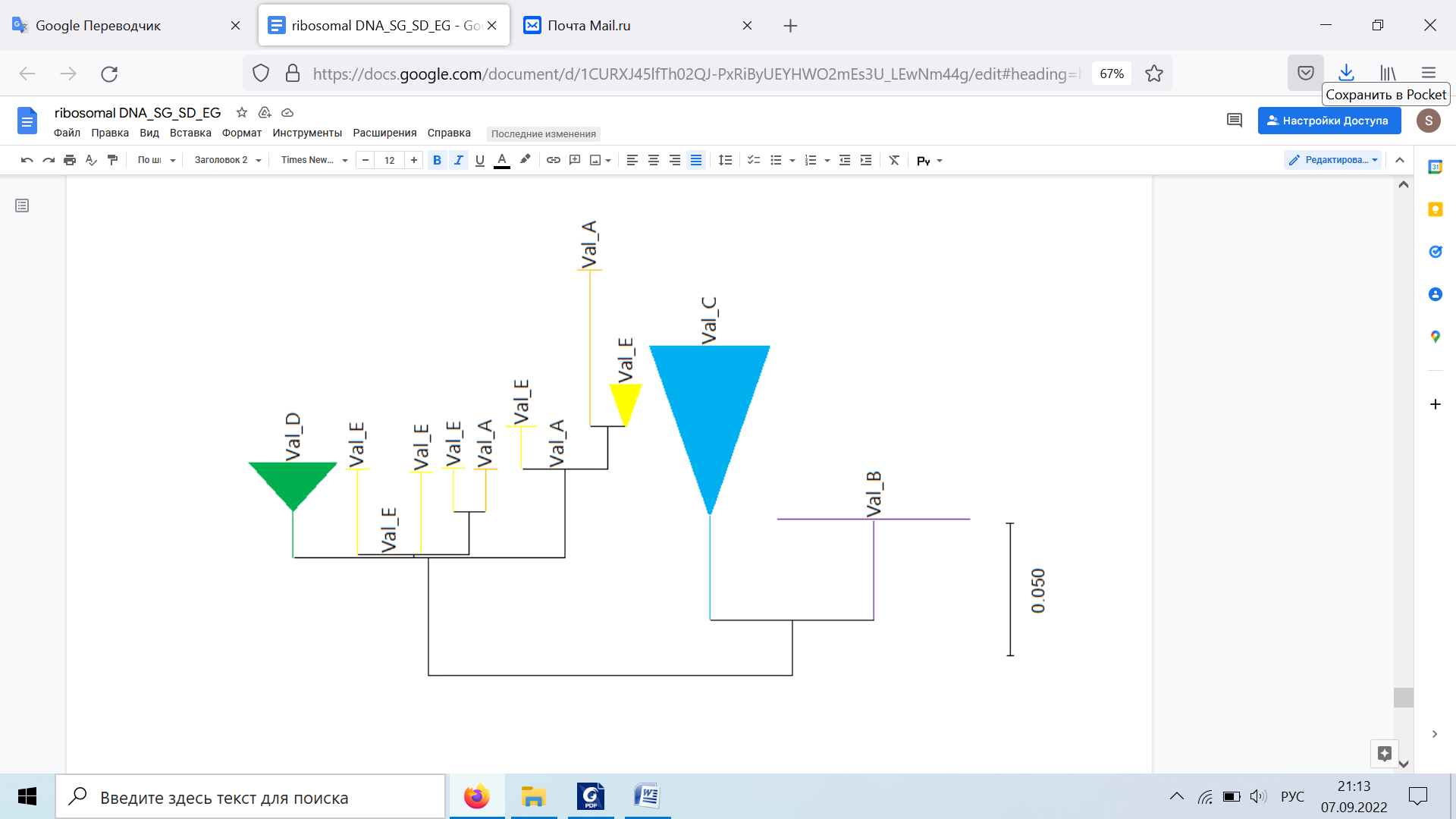
GC% ≈ 76.9

**III. Supplementary Figures**

**Figure S1. Identification of tandem repeat regions in the *N. meleagris* IGS (JABXER010000123) by dot-plot analysis.**

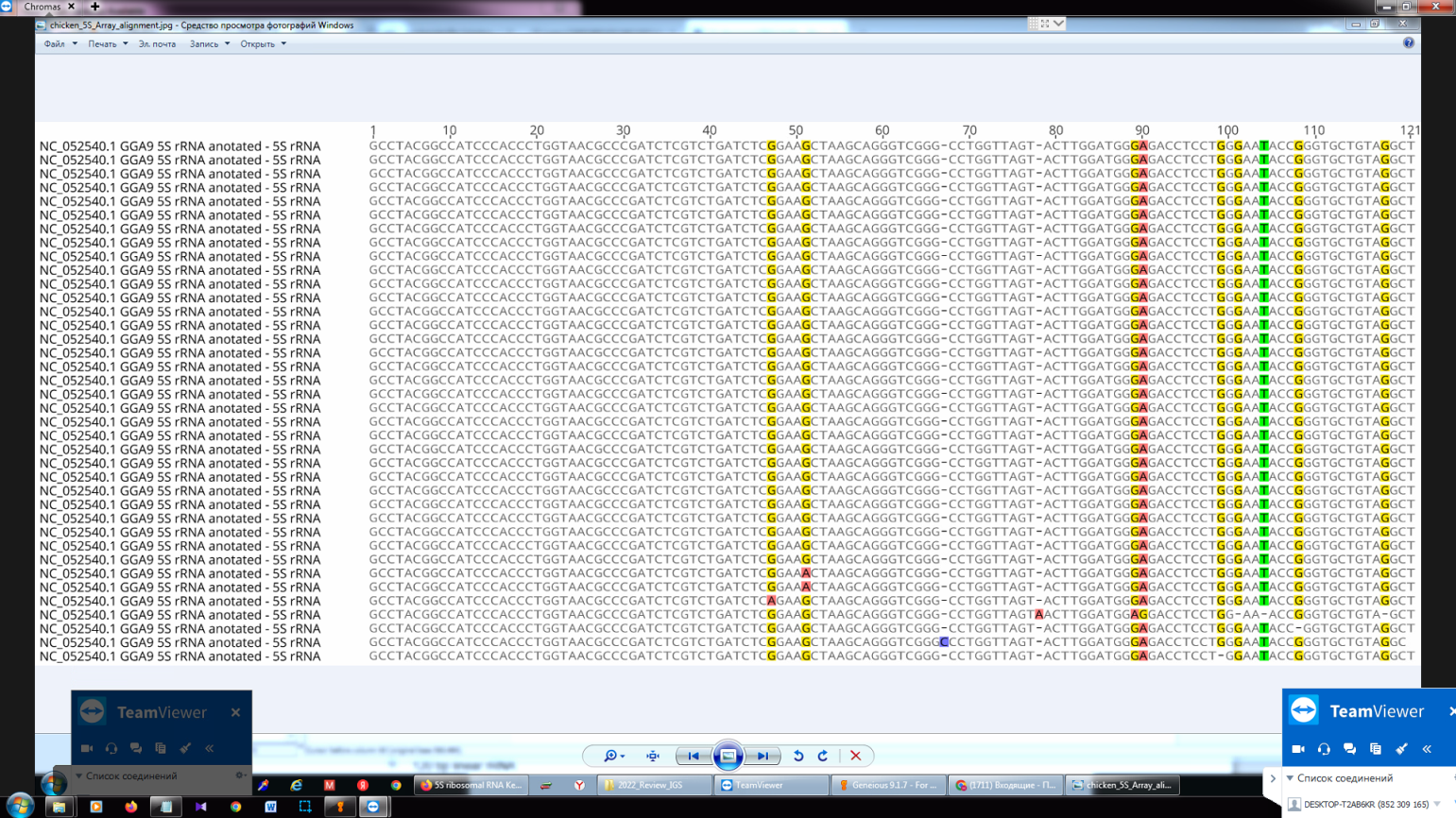


**Figure S2. VAL repeats in *N. meleagris* IGS.** The figure was plotted using the maximum likelihood method. Highly degenerated Val\_F and Val\_FC repeats were not included in the phylogenetic analsis.



**Figure S3. Alignments of the 5S rRNA sequences found in the chicken and guinea fowl genomic sequences. (**a) Alignment of the 5S rRNA genes found organized in an array at ~0.5 Mb on chromosome 9 (NC\_052540.1) in the *G.gallus* reference genome GRCg7w. (b) Alignment of the 5S rRNA gene variants found in chromosome 9 (NC\_052540.1), chromosome 2 (NC\_052533.1), chromosome 6 (NC\_052537.1) in the *G.gallus* reference genome GRCg7w and two types of 5S rRNA downloaded from the NCBI Nucleotide database (Chicken truncated 5S rRNA II type 2, Chicken truncated 5S rRNA I type 1b, Chicken truncated 5S rRNA I type 1a, Chicken 5S rRNA, Chicken 5S rRNA I, Chicken 5S rRNA II). Similarity of the 5S rRNA gene copy from chromosome 2 to rRNA type II are marked by a red frame. (c) Alignment of the 5S rRNA genes found in the JABXER010000004 NCBI WGS contig assembled during *N.meleagris* genome deciphering. Disagreements are marked in colour.

(a)



(b)



(c)

