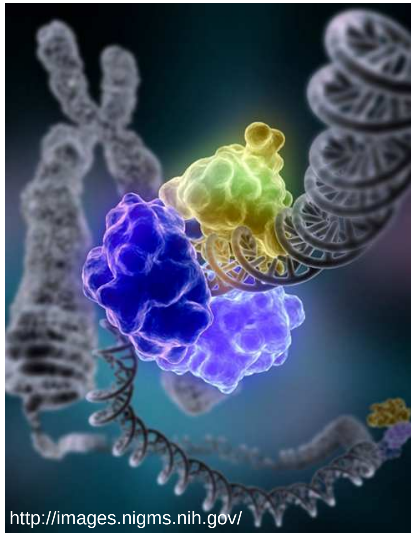
Comprehensive Bio molecular Databases: a worldwide cooperation ?

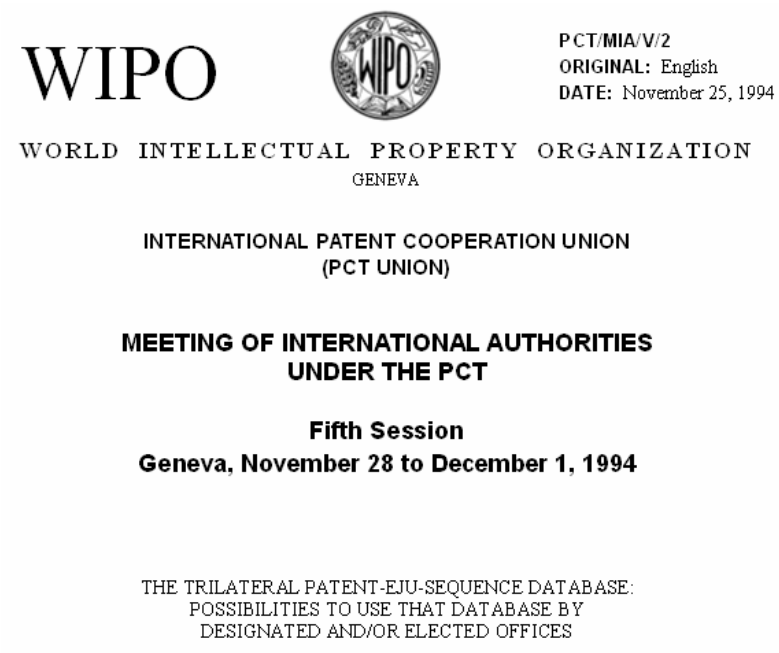

http://images.nigms.nih.gov/

Stéphane Nauche
http://www.linkedin.com/in/snauche

Ana Richart de la Torre

The SLING project is funded by the European Commission within Research Infrastructures of the FP7 Capacities Specific Programme, grant agreement number 226073 (Integrating Activity)

Historical Background



PCT/MIA/V/2
ORIGINAL: English
DATE: November 25, 1994

WIPO

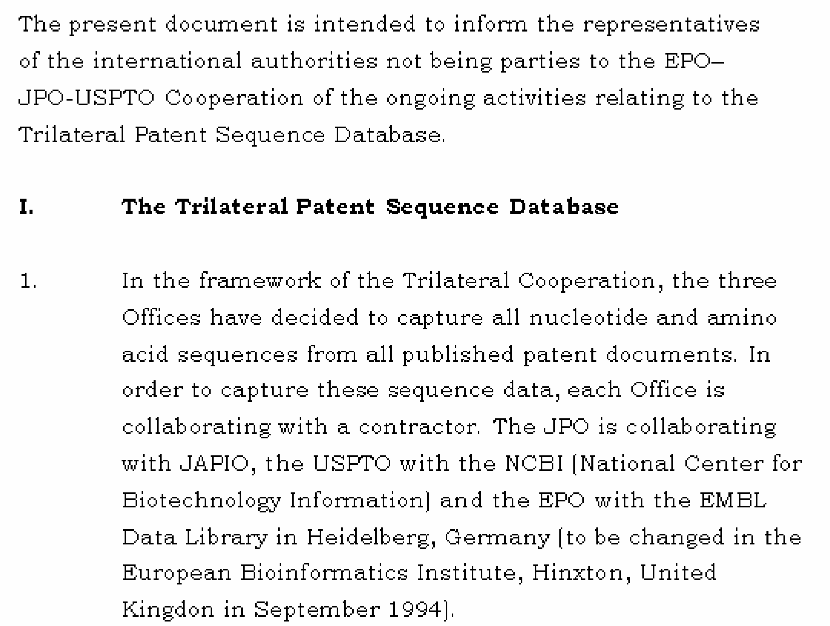
WORLD INTELLECTUAL PROPERTY ORGANIZATION
GENEVA

INTERNATIONAL PATENT COOPERATION UNION
(PCT UNION)

MEETING OF INTERNATIONAL AUTHORITIES
UNDER THE PCT

Fifth Session
Geneva, November 28 to December 1, 1994

THE TRILATERAL PATENT-EJU-SEQUENCE DATABASE:
POSSIBILITIES TO USE THAT DATABASE BY
DESIGNATED AND/OR ELECTED OFFICES

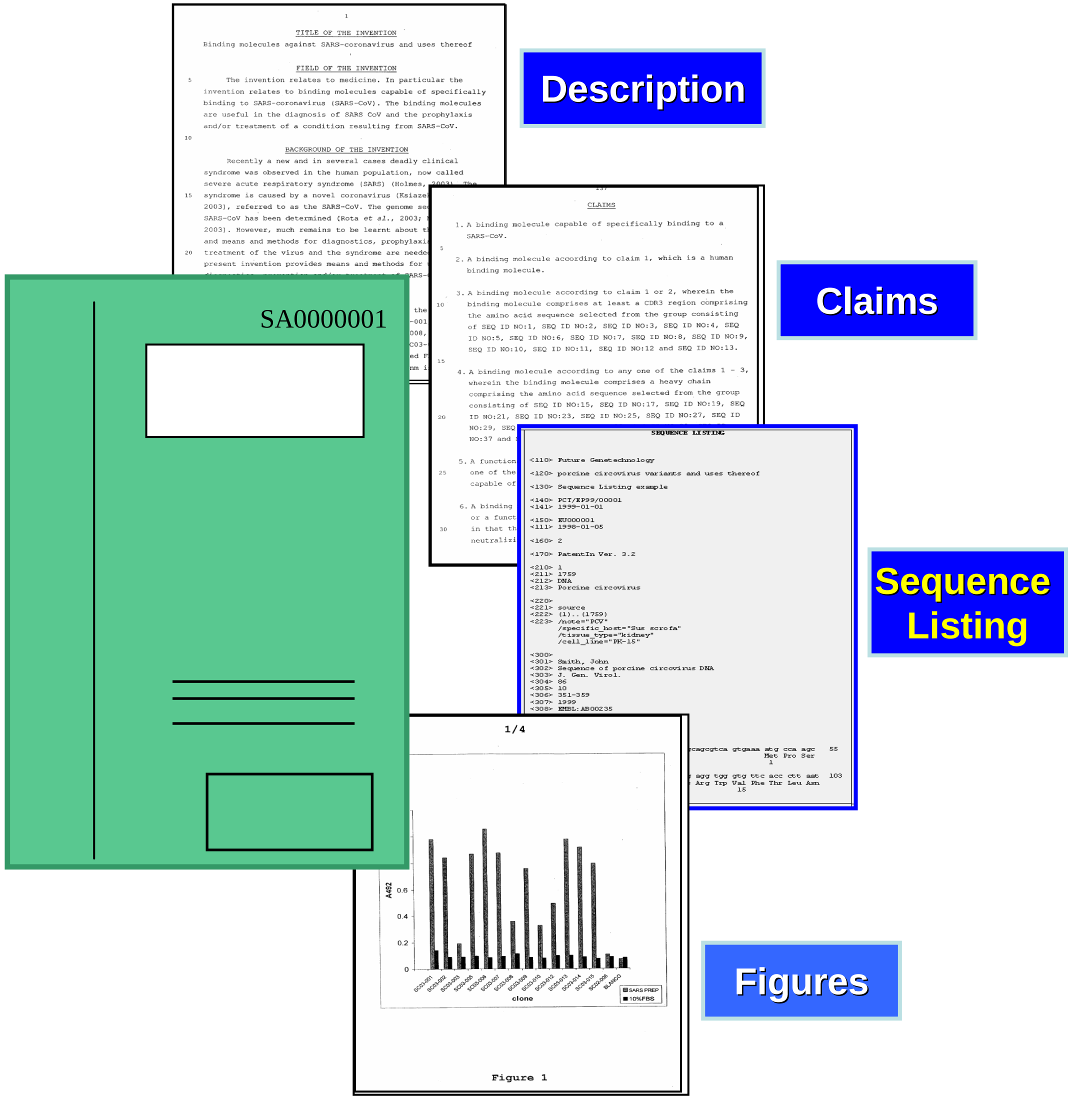


The present document is intended to inform the representatives of the international authorities not being parties to the EPO–JPO-USPTO Cooperation of the ongoing activities relating to the Trilateral Patent Sequence Database.

I. The Trilateral Patent Sequence Database

1. In the framework of the Trilateral Cooperation, the three Offices have decided to capture all nucleotide and amino acid sequences from all published patent documents. In order to capture these sequence data, each Office is collaborating with a contractor. The JPO is collaborating with JAPIO, the USPTO with the NCBI (National Center for Biotechnology Information) and the EPO with the EMBL Data Library in Heidelberg, Germany (to be changed in the European Bioinformatics Institute, Hinxton, United Kingdon in September 1994).

Sequence Listing



1

TITLE OF THE INVENTION

Binding molecules against SARS-coronavirus and uses thereof

FIELD OF THE INVENTION

The invention relates to medicine. In particular the invention relates to binding molecules capable of specifically binding to SARS-coronavirus (SARS-CoV). The binding molecules are useful in the diagnosis of SARS CoV and the prophylaxis and/or treatment of a condition resulting from SARS-CoV.

BACKGROUND OF THE INVENTION

Recently a new and in several cases deadly clinical syndrome was observed in the human population, now called severe acute respiratory syndrome (SARS) (Holmes, 2003). The syndrome is caused by a novel coronavirus (Ksiazek 2003), referred to as the SARS-CoV. The genome se SARS-CoV has been determined (Rota *et al.*, 2003; 2003). However, much remains to be learnt about th and means and methods for diagnostics, prophylaxis treatment of the virus and the syndrome are neede present invention provides means and methods for

the
-001
008,
C03-
ed F
nm i

SA0000001

Description

CLAIMS

1. A binding molecule capable of specifically binding to a SARS-CoV.

2. A binding molecule according to claim 1, which is a human binding molecule.

3. A binding molecule according to claim 1 or 2, wherein the binding molecule comprises at least a CDR3 region comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12 and SEQ ID NO:13.

4. A binding molecule according to any one of the claims 1 - 3, wherein the binding molecule comprises a heavy chain comprising the amino acid sequence selected from the group consisting of SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ NO:37 and

5. A function one of the capable of

6. A binding or a funct in that th neutralizi

Claims

SEQUENCE LISTING

```
<110> Future Genetechnology

<120> porcine circovirus variants and uses thereof

<130> Sequence Listing example

<140> PCT/EP99/00001
<141> 1999-01-01

<150> EU000001
<111> 1998-01-05

<160> 2

<170> PatentIn Ver. 3.2

<210> 1
<211> 1759
<212> DNA
<213> Porcine circovirus

<220>
<221> source
<222> (1)..(1759)
<223> /note="PCV"
      /specific_host="Sus scrofa"
      /tissue_type="kidney"
      /cell_line="PK-15"

<300>
<301> Smith, John
<302> Sequence of porcine circovirus DNA
<303> J. Gen. Virol.
<304> 86
<305> 10
<306> 351-359
<307> 1999
<308> EMBL:AB00235

cagcgtca gtgaaa atg cca agc    55
                Met Pro Ser
                 1

 agg tgg gtg ttc acc ctt aat   103
 Arg Trp Val Phe Thr Leu Asn
            15
```

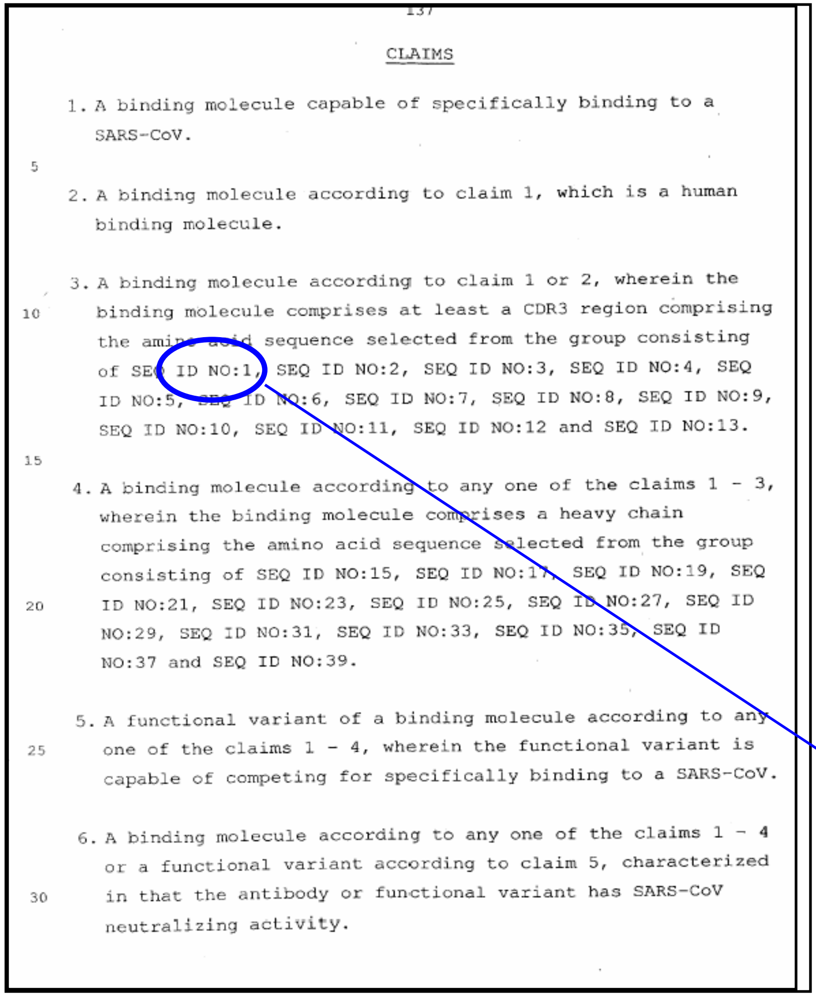
Sequence Listing

1/4

Figure 1

Figures

Sequence Listing



137

CLAIMS

1. A binding molecule capable of specifically binding to a SARS-CoV.

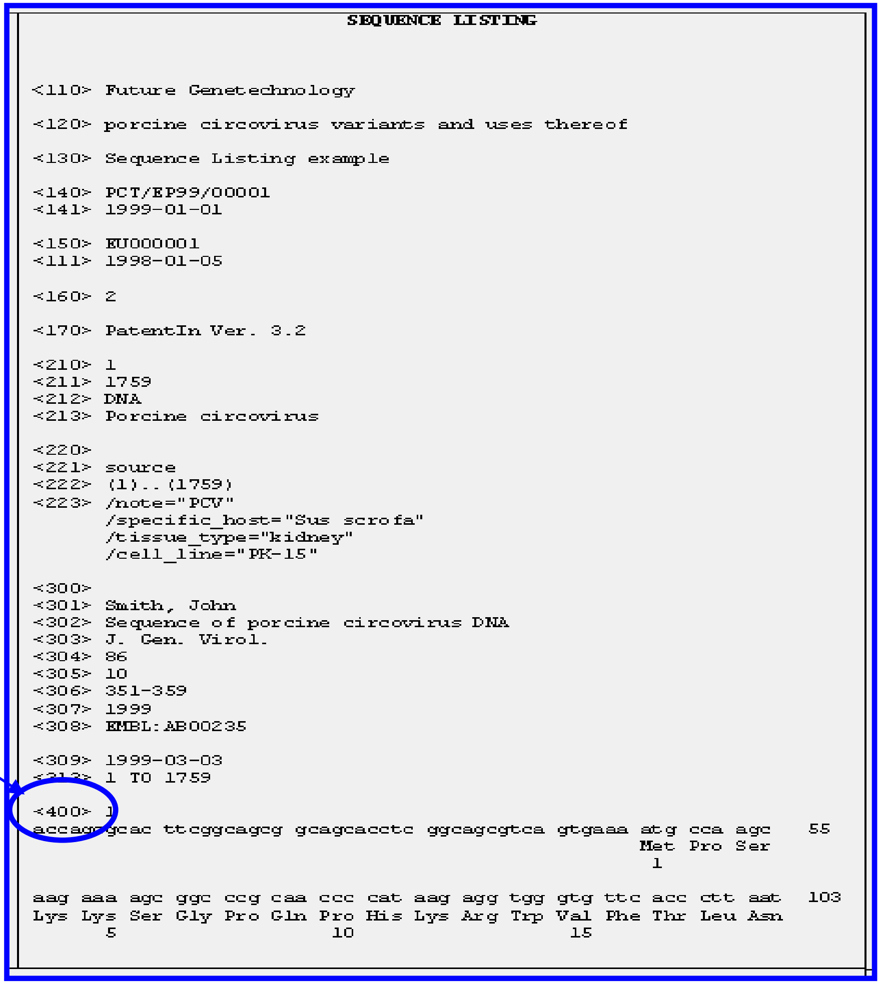
2. A binding molecule according to claim 1, which is a human binding molecule.

3. A binding molecule according to claim 1 or 2, wherein the binding molecule comprises at least a CDR3 region comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12 and SEQ ID NO:13.

4. A binding molecule according to any one of the claims 1 - 3, wherein the binding molecule comprises a heavy chain comprising the amino acid sequence selected from the group consisting of SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37 and SEQ ID NO:39.

5. A functional variant of a binding molecule according to any one of the claims 1 - 4, wherein the functional variant is capable of competing for specifically binding to a SARS-CoV.

6. A binding molecule according to any one of the claims 1 - 4 or a functional variant according to claim 5, characterized in that the antibody or functional variant has SARS-CoV neutralizing activity.

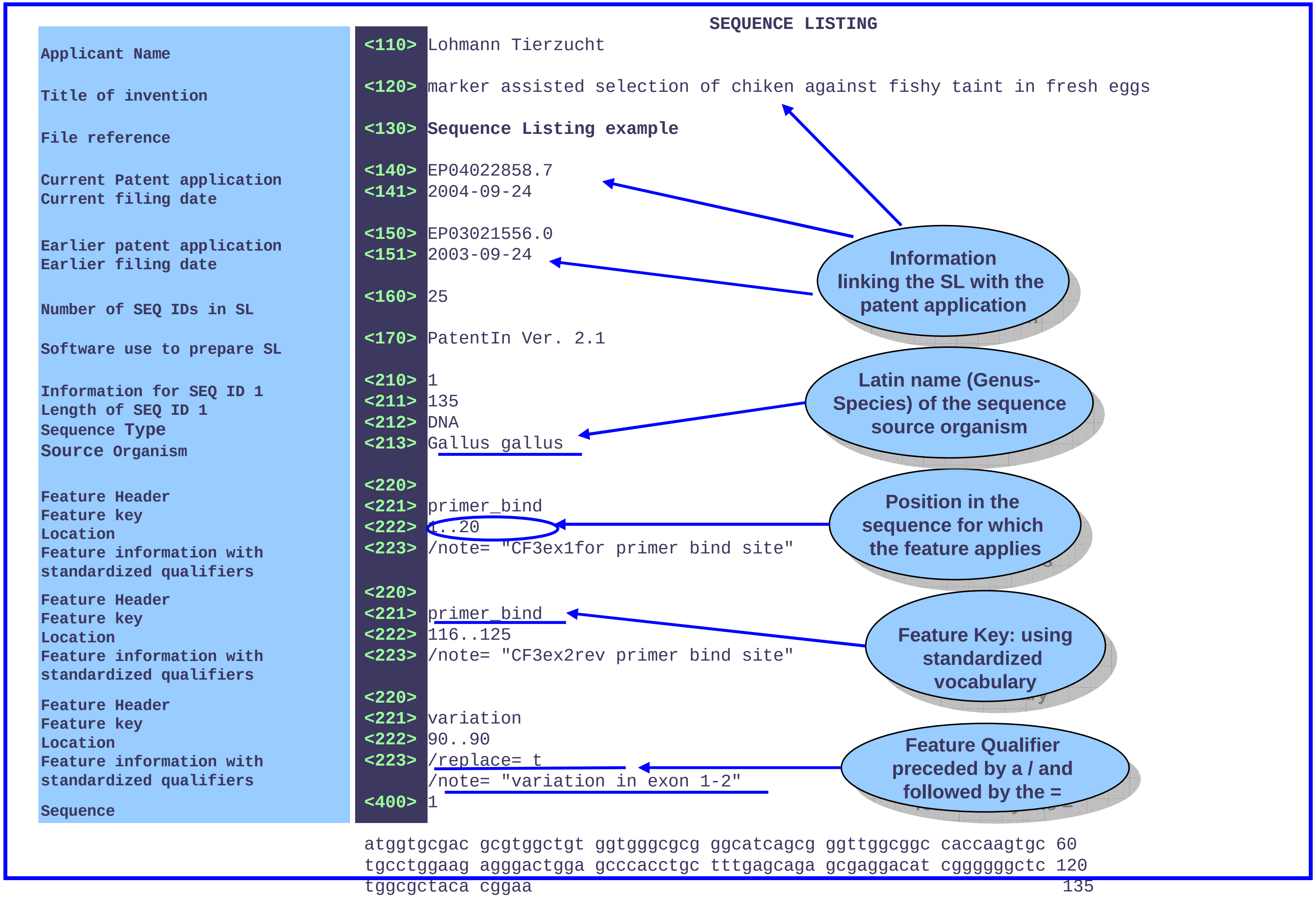


SEQUENCE LISTING

```
<110> Future Genetechnology

<120> porcine circovirus variants and uses thereof

<130> Sequence Listing example

<140> PCT/EP99/00001
<141> 1999-01-01

<150> EU000001
<111> 1998-01-05

<160> 2

<170> PatentIn Ver. 3.2

<210> 1
<211> 1759
<212> DNA
<213> Porcine circovirus

<220>
<221> source
<222> (1)..(1759)
<223> /note="PCV"
      /specific_host="Sus scrofa"
      /tissue_type="kidney"
      /cell_line="PK-15"

<300>
<301> Smith, John
<302> Sequence of porcine circovirus DNA
<303> J. Gen. Virol.
<304> 86
<305> 10
<306> 351-359
<307> 1999
<308> EMBL:AB00235

<309> 1999-03-03
<313> 1 TO 1759

<400> 1
accag[illegible]gcac ttcggcagcg gcagcacctc ggcagcgtca gtgaaa atg cca agc   55
                                                          Met Pro Ser
                                                           1

aag aaa agc ggc ccg caa ccc cat aag agg tgg gtg ttc acc ctt aat   103
Lys Lys Ser Gly Pro Gln Pro His Lys Arg Trp Val Phe Thr Leu Asn
     5                  10                  15
```

SLING
Serving Life-science Information for the Next Generation

WIPO ST 25



SEQUENCE LISTING

Applicant Name	<110>	Lohmann Tierzucht
Title of invention	<120>	marker assisted selection of chiken against fishy taint in fresh eggs
File reference	<130>	**Sequence Listing example**
Current Patent application	<140>	EP04022858.7
Current filing date	<141>	2004-09-24
Earlier patent application	<150>	EP03021556.0
Earlier filing date	<151>	2003-09-24
Number of SEQ IDs in SL	<160>	25
Software use to prepare SL	<170>	PatentIn Ver. 2.1
Information for SEQ ID 1	<210>	1
Length of SEQ ID 1	<211>	135
Sequence Type	<212>	DNA
Source Organism	<213>	Gallus gallus
Feature Header	<220>	
Feature key	<221>	primer_bind
Location	<222>	1..20
Feature information with standardized qualifiers	<223>	/note= "CF3ex1for primer bind site"
Feature Header	<220>	
Feature key	<221>	primer_bind
Location	<222>	116..125
Feature information with standardized qualifiers	<223>	/note= "CF3ex2rev primer bind site"
Feature Header	<220>	
Feature key	<221>	variation
Location	<222>	90..90
Feature information with standardized qualifiers	<223>	/replace= t /note= "variation in exon 1-2"
Sequence	<400>	1

```
atggtgcgac gcgtggctgt ggtgggcgcg ggcatcagcg ggttggcggc caccaagtgc 60
tgcctggaag agggactgga gcccacctgc tttgagcaga gcgaggacat cggggggctc 120
tggcgctaca cggaa                                                  135
```

Information linking the SL with the patent application

Latin name (Genus-Species) of the sequence source organism

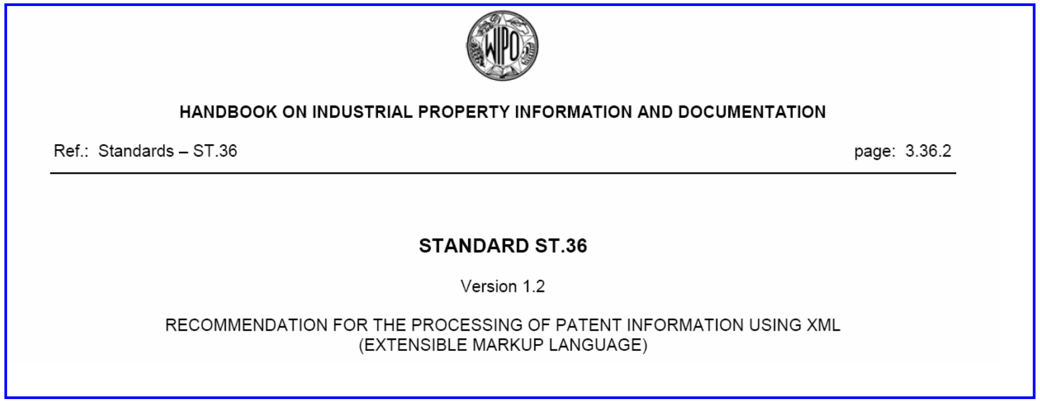
Position in the sequence for which the feature applies

Feature Key: using standardized vocabulary

Feature Qualifier preceded by a / and followed by the =

XML Sequence Listing

General Information



HANDBOOK ON INDUSTRIAL PROPERTY INFORMATION AND DOCUMENTATION

Ref.: Standards – ST.36 page: 3.36.2

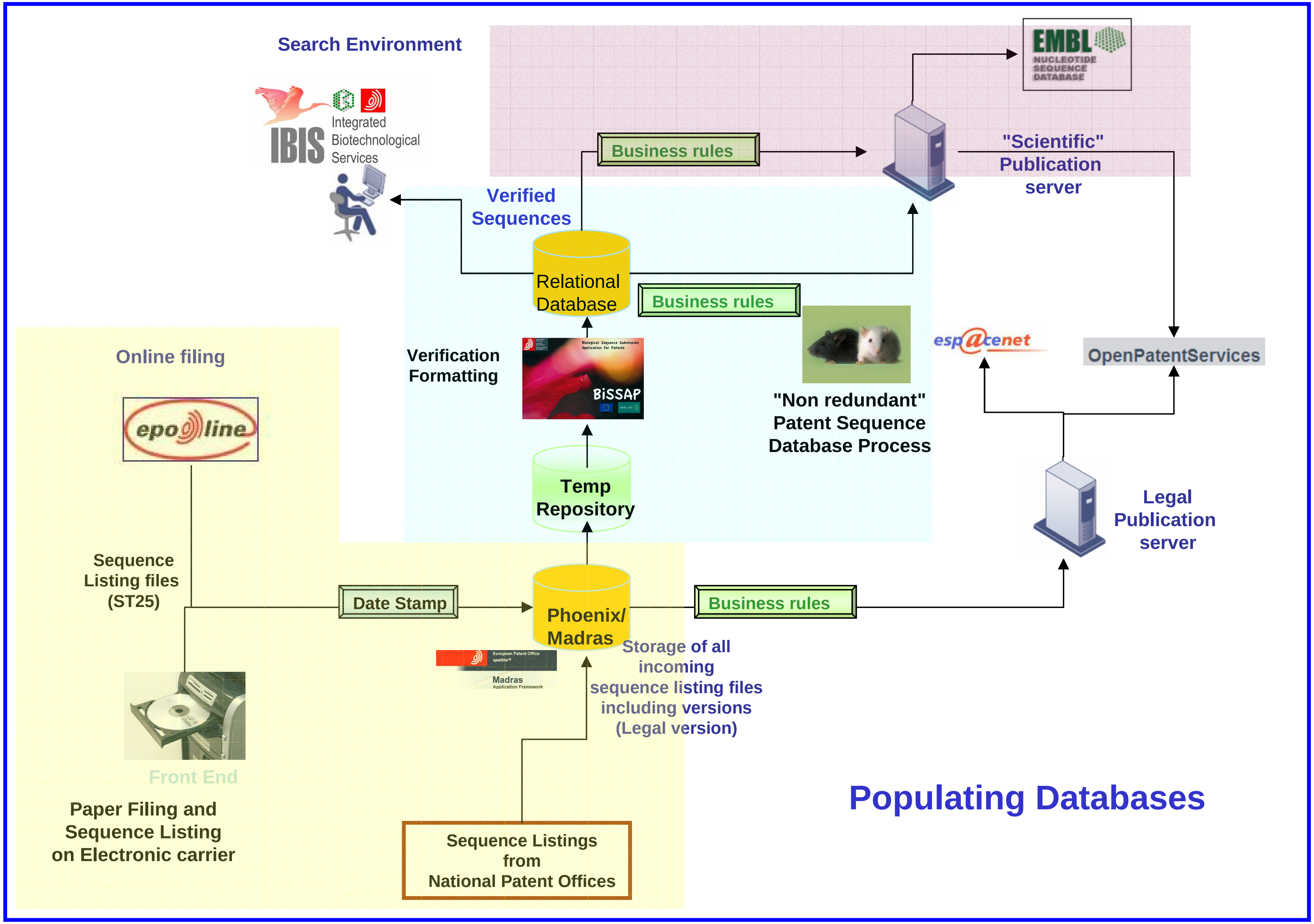
STANDARD ST.36

Version 1.2

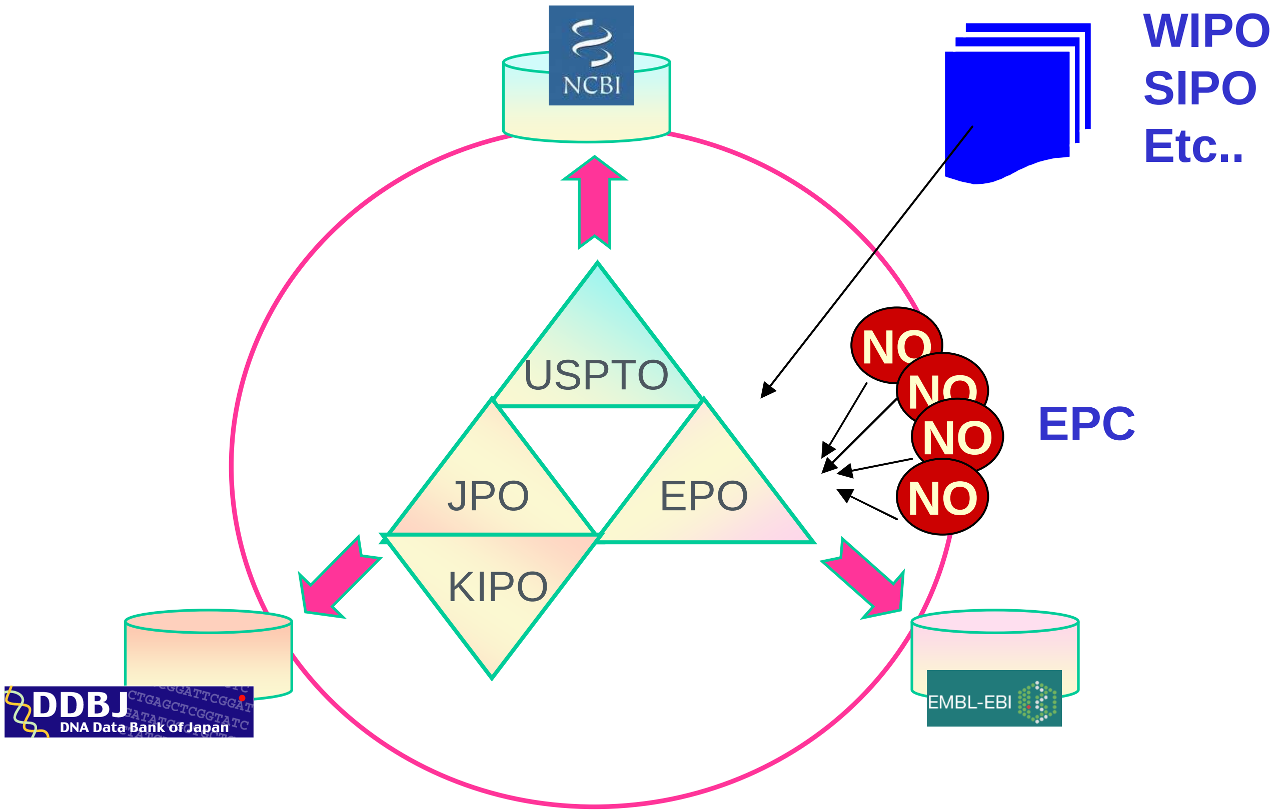
RECOMMENDATION FOR THE PROCESSING OF PATENT INFORMATION USING XML
(EXTENSIBLE MARKUP LANGUAGE)

Sequence Part

Populating Databases



International Cooperation



NCBI

WIPO
SIPO
Etc..

USPTO

JPO

EPO

KIPO

NO
NO
NO
NO

EPC

DDBJ
DNA Data Bank of Japan

EMBL-EBI

Necessity to increase coverage and avoid redundant Information

- Increase in Biotechnology applications disclosing biological sequences.

To date, only **JPO, USPTO, EPO** and **KIPO**, send their sequences to the public repositories (NCBI/DDBJ/EBI)

- Data spreading from other patent offices is not included in the exchange flow.

- June 2003: completeness study performed by the EPO.

  (NCBI/DDBJ/EBI) vs. GeneSeq = Differences in coverage were detected :none is exhaustive !

- April 2005 (Washington): Project initiation report presented by the EPO to TO's

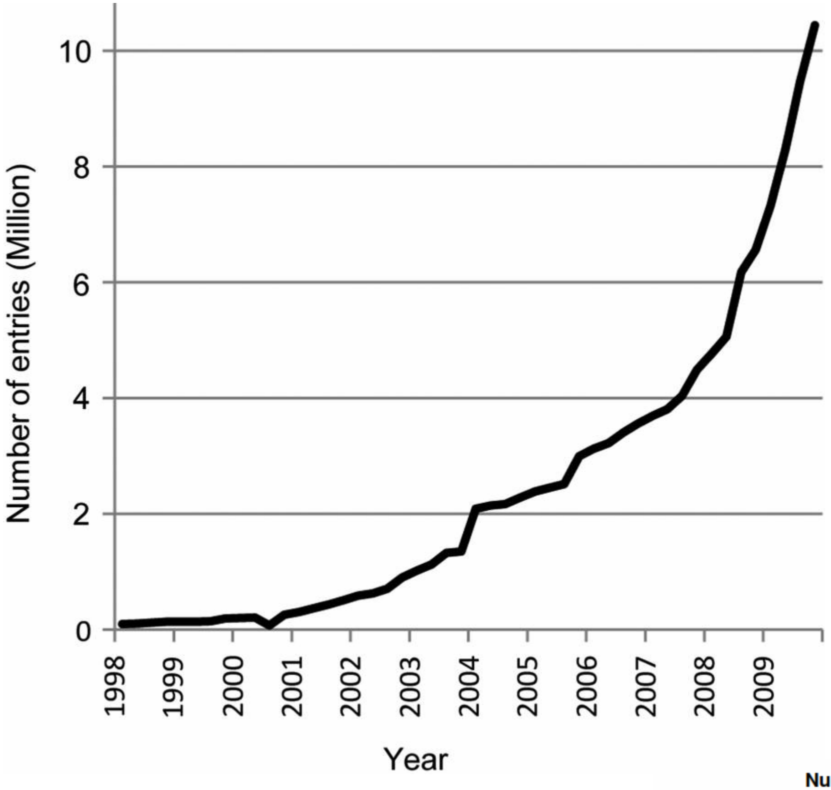
  Aims: generation of Non-redundant Patent sequence databases with increased coverage

- (Tokyo, April 2007, Munich, November 2007 and Washington 2009):

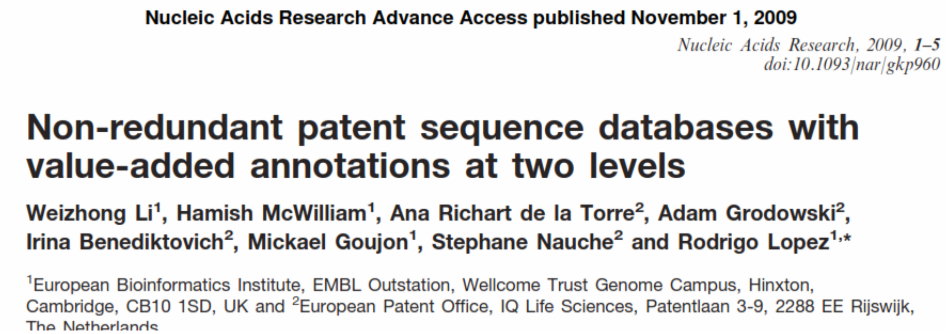
➢ *EPO progresses were presented to the TO's*
➢ *JPO involved KIPO in the data exchange flow to the public sequence repositories*

EPO and EBI have completed the generation of Non-redundant Patent sequence Databases, which are publicly available through the EBI services

Patent Sequence Databases Figures



August 2010 :
17 million entries



Nucleic Acids Research Advance Access published November 1, 2009

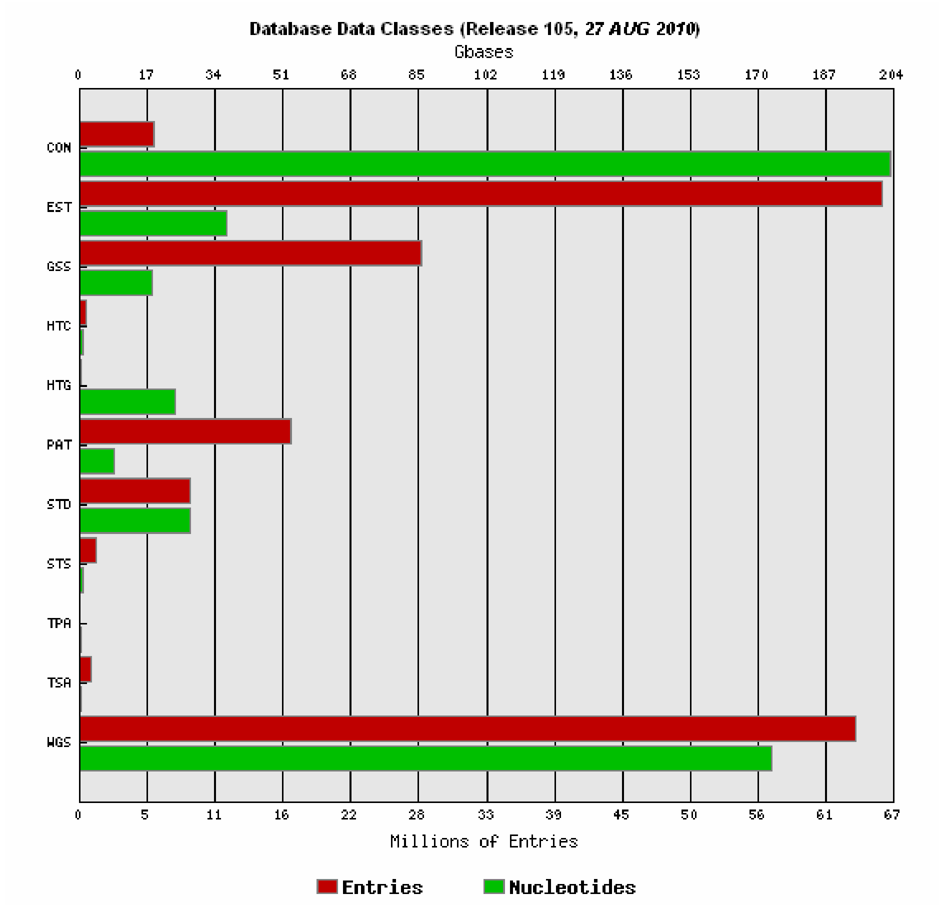
Nucleic Acids Research, 2009, ***1–5***
doi:10.1093/nar/gkp960

Non-redundant patent sequence databases with value-added annotations at two levels

Weizhong Li[1], Hamish McWilliam[1], Ana Richart de la Torre[2], Adam Grodowski[2], Irina Benediktovich[2], Mickael Goujon[1], Stephane Nauche[2] and Rodrigo Lopez[1,*]

[1]European Bioinformatics Institute, EMBL Outstation, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK and [2]European Patent Office, IQ Life Sciences, Patentlaan 3-9, 2288 EE Rijswijk, The Netherlands

Patent Sequence Databases Figures



- •PRT sequences sent to EMBL - 224196

  4.53% increase of EMBL patent PRT db

- •DNA sequences sent to EMBL - 498189

  2.84% increase of EMBL patent DNA db

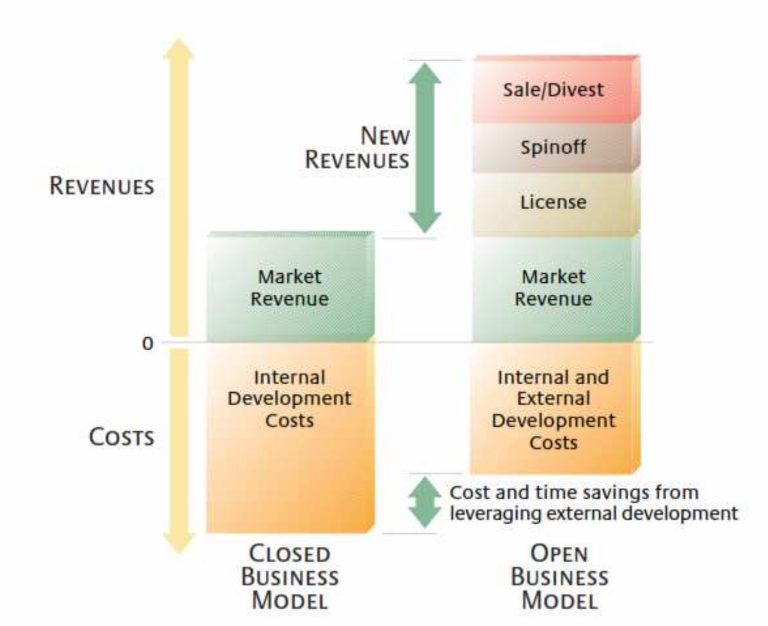
Open Innovation

Connect and Develop

Inside Procter & Gamble's New Model for Innovation

Procter & Gamble's radical strategy of open innovation now produces more than 35% of the company's innovations and billions of dollars in revenue.

Harvard Business Review, 2004, Huston et al.“



MIT Sloan Management Review; 2007; Chesbrough h; „Why Companies should have open Business Models“

yourencore®
ACCELERATING INNOVATION THROUGH PROVEN EXPERIENCE

PRO INNO EUROPE
INNO ACTIONS
ip4inno

Summary

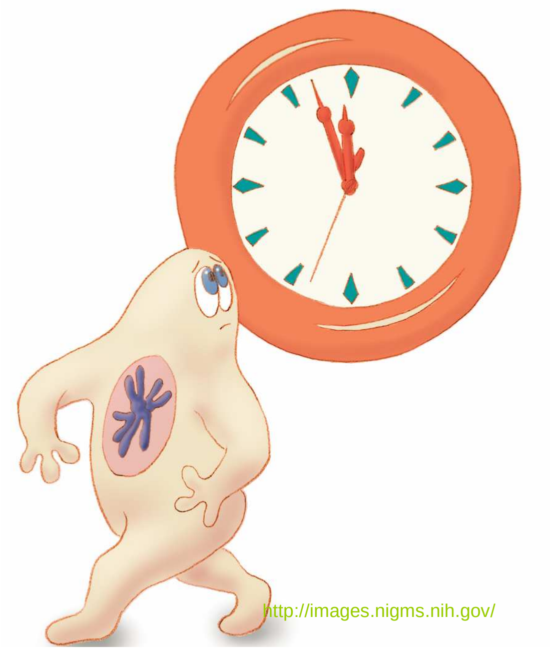
Create sequences

Collect

Populate Patent sequence databases

Disseminate Information
Search ‚Prior art‘ in Biotechnology

Questions



http://images.nigms.nih.gov/

Thank
You