Biobanking for Molecular Epidemiology:
The IARC experience

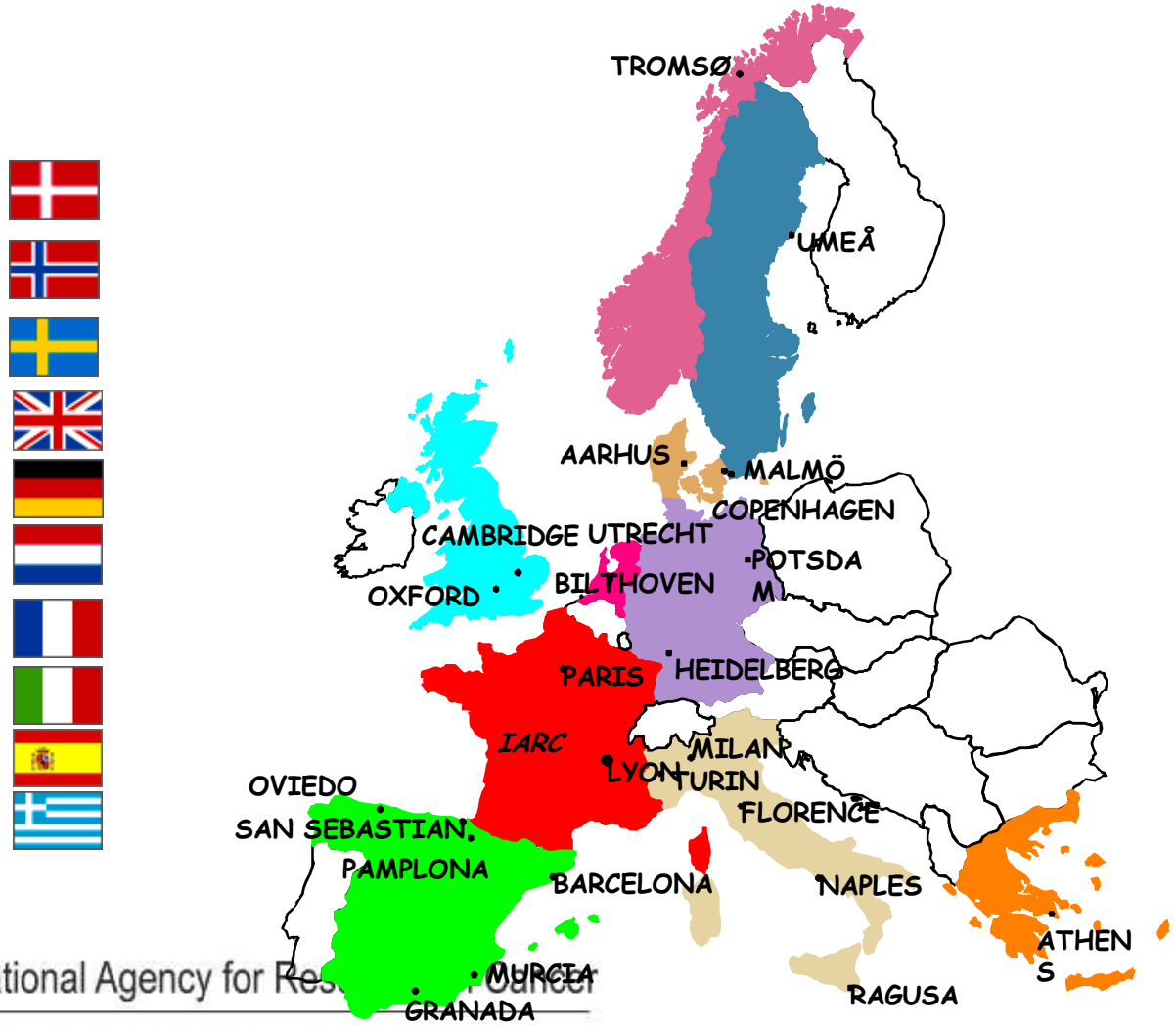
Pierre Hainaut
International Agency for Research on Cancer
World Health Organization

International Agency for Research on Cancer
Centre International de Recherche sur le Cancer

Outline

- Example of large Epidemiology study and biobank: EPIC
- Example of large Genetic Epidemiology study: lung cancer, Hung et al. Nature 2008; McKay et al. Nature Genetics in press
- Flexible biobanking solutions for flexible research networks: The International Liver Cancer study
- An integrated approach for biobank networks

European Prospective Investigation into Cancer (EPIC)



International Agency for Research on Cancer
Centre International de Recherche sur le Cancer

EPIC: European Prospective Investigation into Cancer

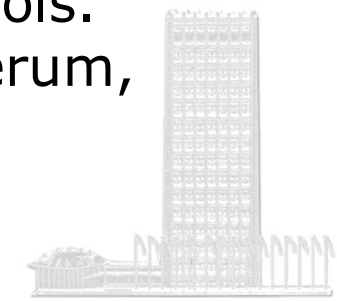
Multi-centre prospective study aimed at investigating the relations between nutrition, lifestyle factors and aetiology of cancer and other chronic diseases.

Initiated in 1993 with the collection of data and blood samples in 22 regional centres located in 9 European countries.

Field work completed in 1998 with the inclusion of 521,483 subjects. All had provided questionnaire data.

395,713 blood samples collected and stored under LN2 vapour for future analyses on cancer cases and controls. This represents over 7 millions aliquots of plasma, serum, WBC, RBC

International Agency for Research on Cancer
Centre International de Recherche sur le Cancer

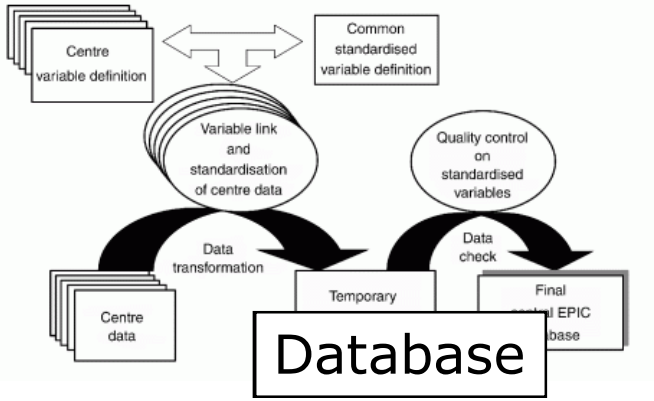


EPIC work scheme

General population

COHORT

Cancer Registration



Database

Active Follow-up

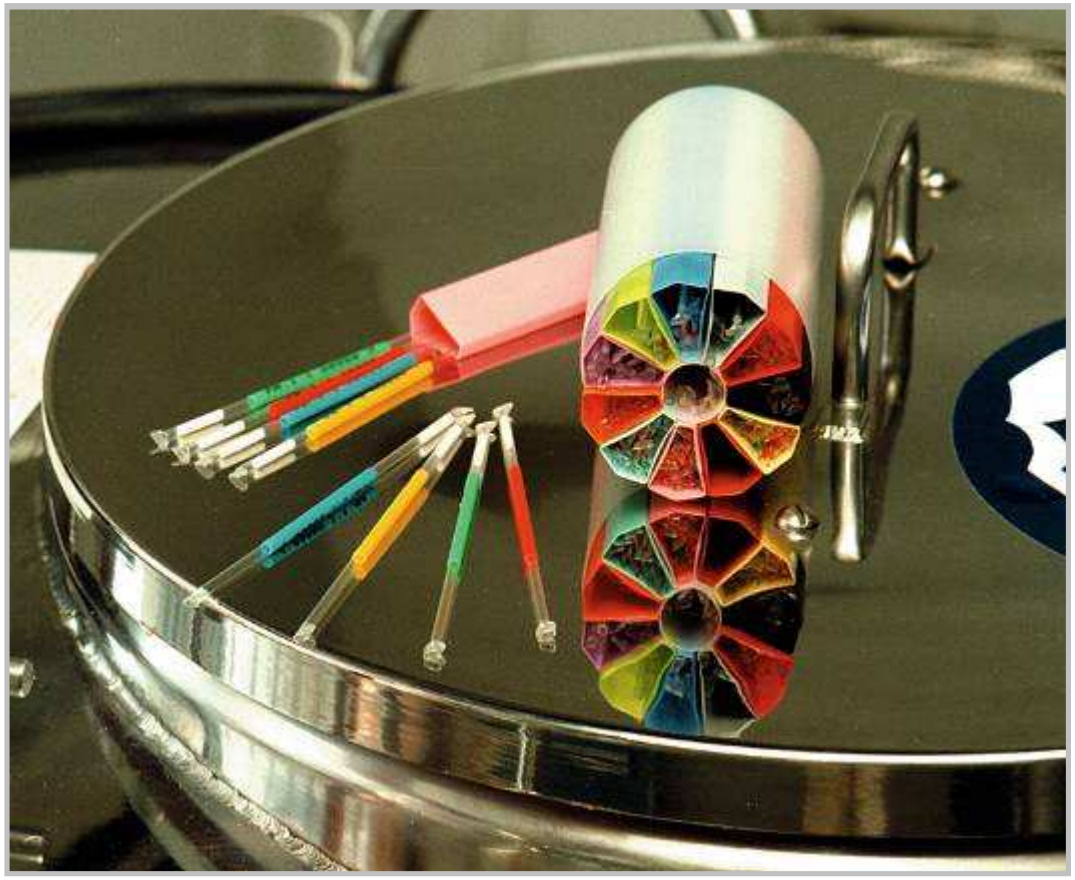
Blood sample

Disease end-points

International Agency for Research on Cancer

Centre International de Recherche sur le Cancer

EPIC Biorepository: storage



30 ml blood:

Plasma: 12 x 500 µl (red straws)

Serum: 8x 500 µl (yellow straws)

Buffy coat: 4x 500 µl (blue straws)

Red blood cells: 4 x 500 µl (green straws)

Total: 24 x 500 µl 300.000 subjects = 7.2 Million aliquots

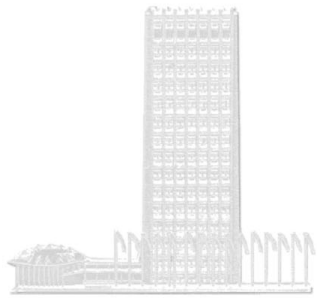
International Agency for Research on Cancer
Centre International de Recherche sur le Cancer

EPIC: incident cancer cases

											Total
Breast	25	485	281	537	199	324	2048	350	187	22	4458
Colo-rectum	25	244	122	289	96	88	179	93	89	13	1238
Prostate	-	391	44	218	119	3	-	33	30	3	841
Stomach	1	52	10	49	24	17	11	47	25	7	243
Lung	14	158	114	181	74	45	94	55	41	14	790

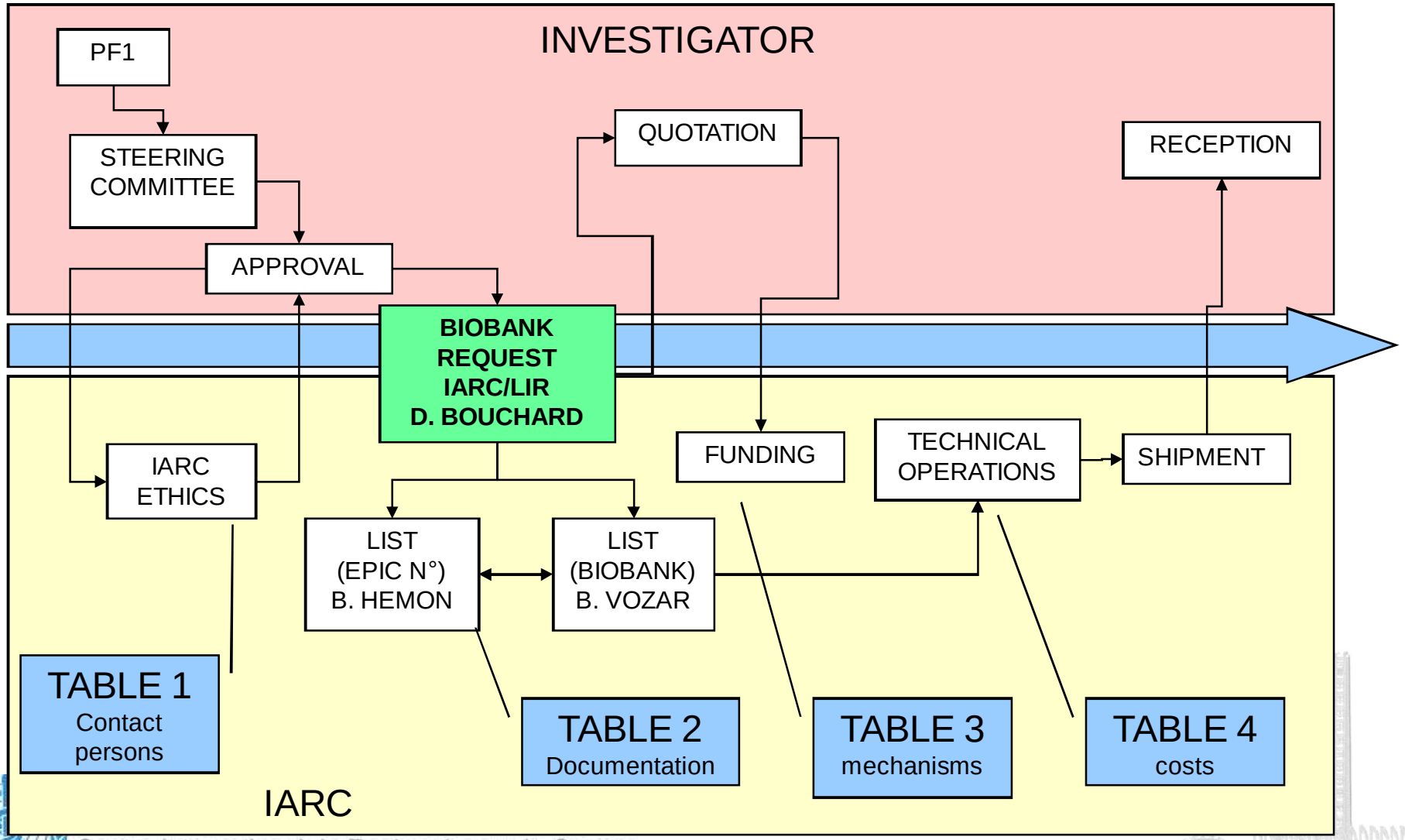
International Agency for Research on Cancer

Centre International de Recherche sur le Cancer



EPIC BRC at IARC: storage, processing, shipping and analysis

EPIC biobank flow chart



Costs of routine technical operations

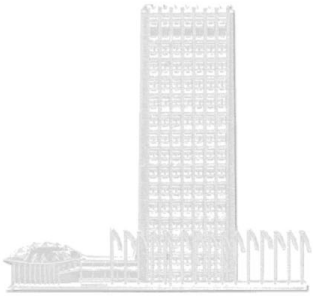
SPECIMEN RETRIEVAL from the EPIC BIOBANK at IARC: 9.00 Euros/subject

Cost per retrieved set of specimens from the EPIC LN2 tanks (One or multiple straws of the same subject)

Salary (basis: LY4 standard staff costs, 1100 specimens retrieved / month / person)	6.20
Liquid Nitrogen	0.50
Consumables	0.50
Packaging, dry ice, temp.storage & shipment	0.70
Sub total per subject	7.90
Overheads 13%	1.03
Total per subject	**8.93**
	(Rounded up to 9.00 Euros)

DNA EXTRACTION:	**12.45 Euros/sample**
Salary	3.40
Consumables	8.05
Overheads 13%	
Total	**12.45 Euros**

ALIQUOTING: **2.00 Euros per sample (incl Overheads)**



Large scale DNA extraction

<u>Pass rate for GWAS (n=4500)</u>

1 straw: 86.4%

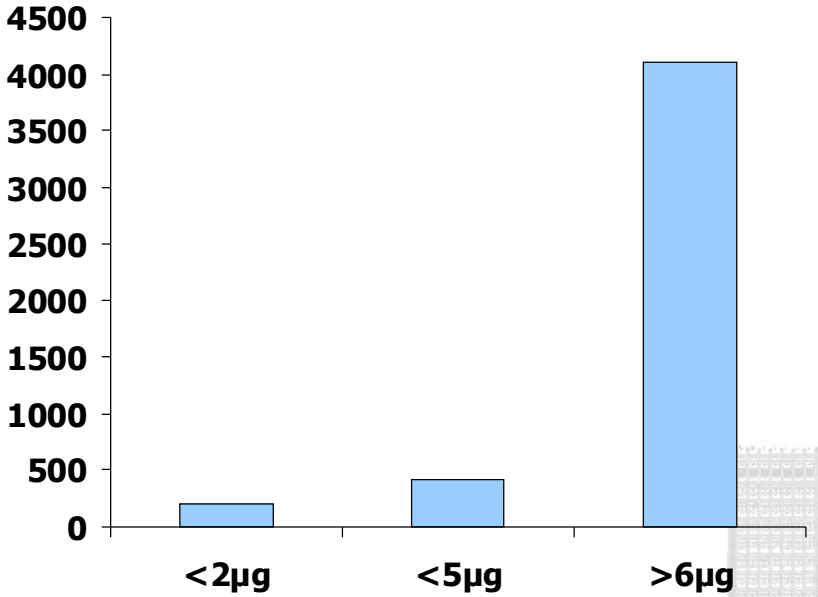
2 straws: 96.4%

<u>Average DNA yield:</u>

24.9±10 µg (n=13562)

<u>Average DNA concentration</u>:

114±64 ng/ml (n=13652)



Genome Wide Association Studies for Lung Cancer

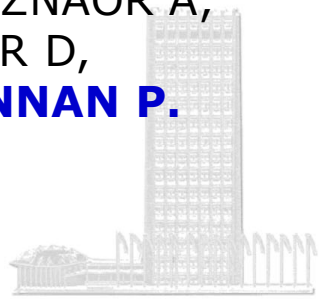
A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25.

Nature, **452(7187):** 633-7, 2008

HUNG RJ, MCKAY JD, GABORIEAU V, BOFFETTA P, HASHIBE M, ZARIDZE D, MUKERIA A, SZESZENIA-DABROWSKA N, LISSOWSKA J, RUDNAI P, FABIANOVA E, MATES D, BENCKO V, FORETOVA L, JANOUT V, CHEN C, GOODMAN G, FIELD JK, LILOGLOU T, XINARIANOS G, CASSIDY A, MCLAUGHLIN J, LIU G, NAROD S, KROKAN HE, SKORPEN F, ELVESTAD MB, HVEEM K, VATTEN L, LINSEISEN J, CLAVEL-CHAPELON F, VINEIS P, BUENO-DE-MESQUITA HB, LUND E, MARTINEZ C, BINGHAM S, RASMUSON T, **HAINAUT P,** RIBOLI E, AHRENS W, BENHAMOU S, LAGIOU P, TRICHOPOULOS D, HOLCÁTOVÁ I, MERLETTI F, KJAERHEIM K, AGUDO A, MACFARLANE G, TALAMINI R, SIMONATO L, LOWRY R, CONWAY DI, ZNAOR A, HEALY C, ZELENIKA D, BOLAND A, DELEPINE M, FOGLIO M, LECHNER D, MATSUDA F, BLANCHE H, GUT I, HEATH S, **LATHROP M, and BRENNAN P.**

International Agency for Research on Cancer

Centre International de Recherche sur le Cancer



East-European Lung Cancer Study

International Agency for Research on Cancer

Centre International de Recherche sur le Cancer

Lung cancer samples

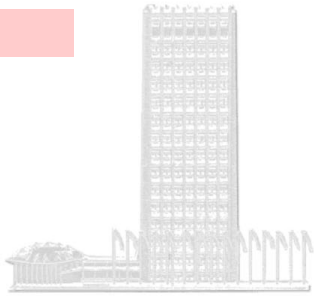
Study	Case	Controls	Countries of Origin
Genome-Wide Association			
Central Europe[1]	1 841	2 441	6 Eastern European countries
Toronto[1]	330	500	Canada
HUNT2/TromsØ[2]	403	412	Norway
CARET[3]	397	392	USA
Total	2 971	3 745	
Replication			
EPIC[2]	1 213	2 591	10 Western European countries
Szczecin[1]	908	1 037	Poland
CARET2[2]	363	1 128	USA
Liverpool[1]	415	817	UK
Total	2 899	5 573	
Overall Total	5 870	9 318	

1st GWA study

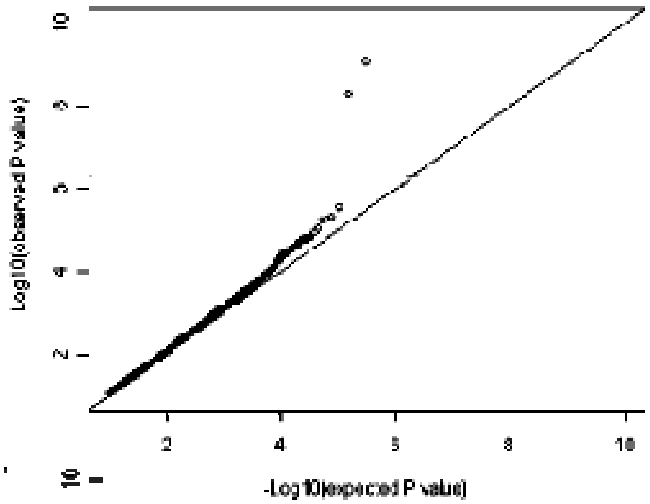
[1] Case-Control collection [2] Cohort collection

International Agency for Research on Cancer

Centre International de Recherche sur le Cancer



Lung cancer 1st phase GWA scan

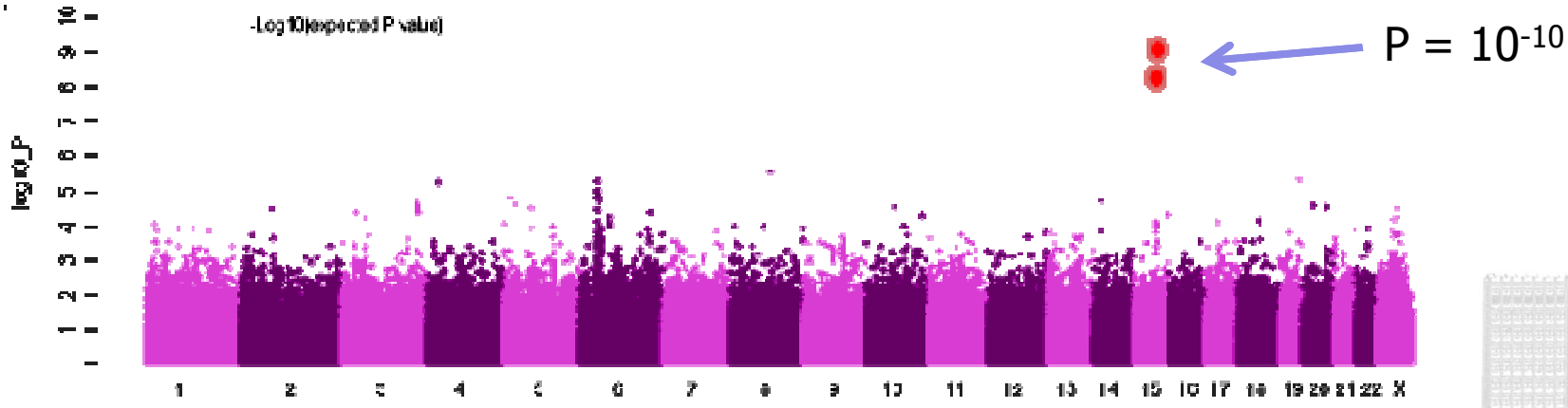


1989 lung cancer cases

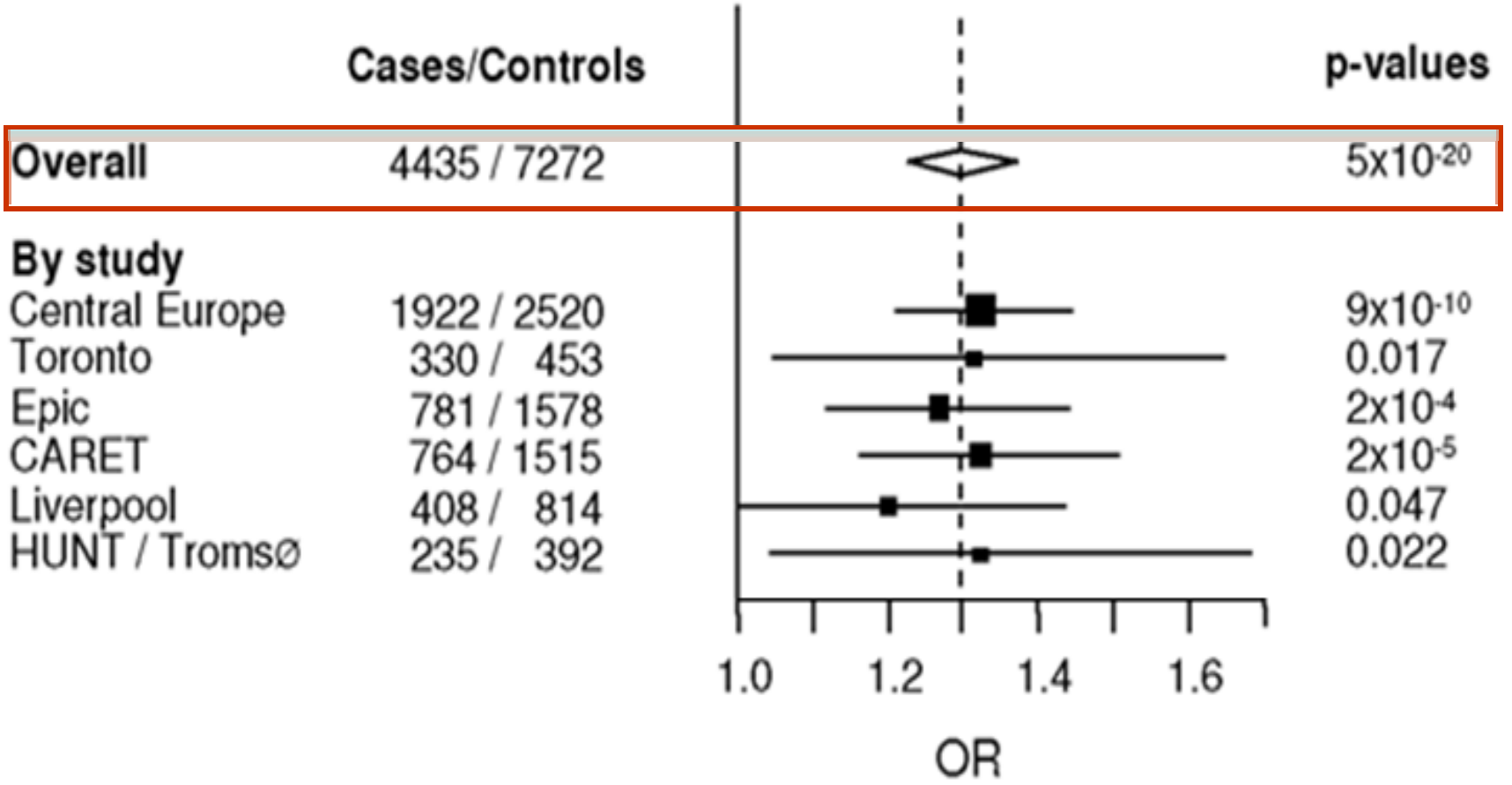
2625 hospital matched controls

310,023 SNPs

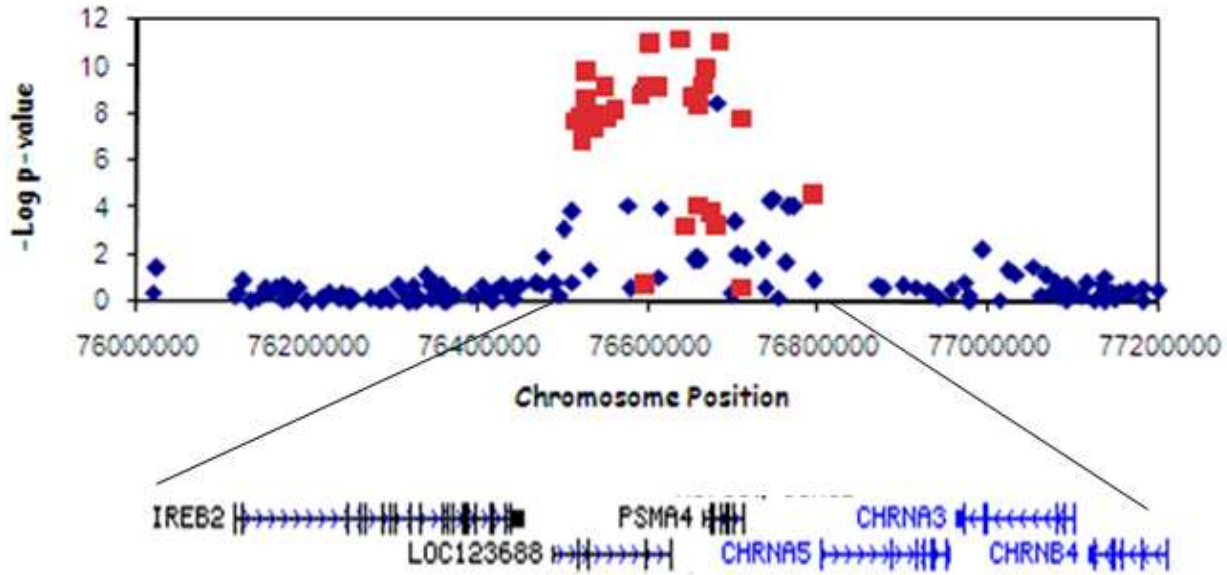
Genome-wide significance $p<5x10^{-7}$



Consistent replication of lung cancer association

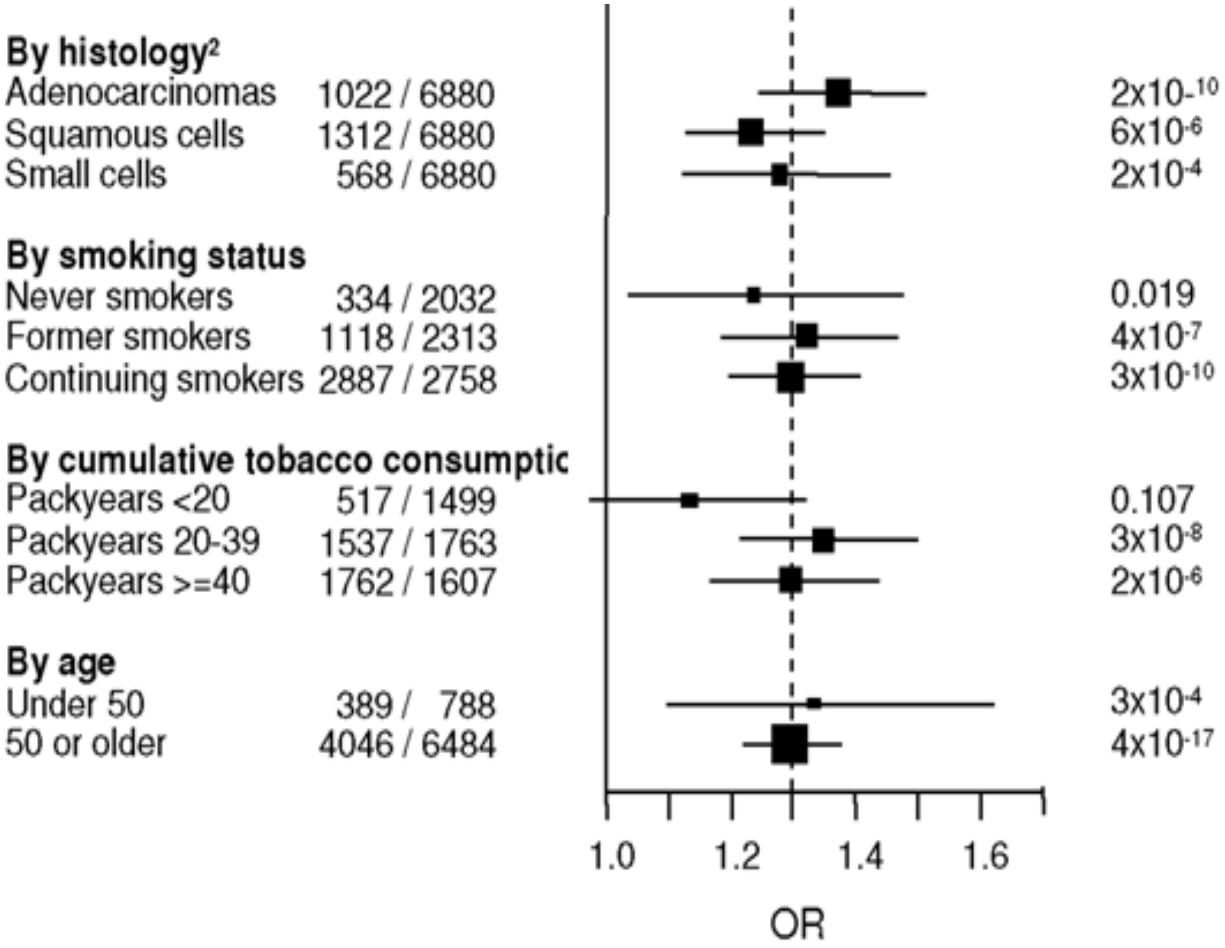


Lung cancer association region



CHRNA3,CHRNA5,CHRNB4
Nicotinic acetylcholine receptor subunits expressed in multiple cell types
Implicated in nicotine dependence
Bind to nicotine and potent lung carcinogens

No association with histology, smoking, age



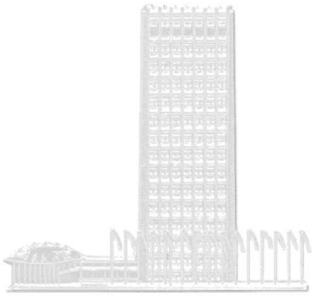
No association with tobacco dependence

	rs8034191 T/T		C/T		C/C		Chi-square p-values
	N	%	N	%	N	%	
Total number of controls	366	100.0	446	100.0	136	100.0	
Smoking status							
Former smokers	195	53.3	243	54.5	71	52.2	0.879
Current smokers	171	46.7	203	45.5	65	47.8	
Duration of smoking (in years)							
1 -9	30	8.2	49	11.0	14	10.3	0.869
10-19	61	16.6	63	14.2	19	14.0	
20-29	84	23.1	93	20.9	32	23.5	
30-39	100	27.5	129	29.0	41	30.2	
>=40	89	24.5	111	24.9	30	22.1	
Number of cigarettes per day							
<10	85	23.4	98	22.0	33	24.3	0.714
10-19	145	39.6	185	41.6	55	40.4	
20-29	82	22.5	87	19.6	32	23.5	
>=30	52	14.3	75	16.9	16	11.8	
First cigarette after waking up							
>31 minutes	56	41.8	74	46.5	25	47.2	0.808
6-30 minutes	52	38.8	51	32.1	18	34.0	
Within 5 minutes	26	19.4	34	21.4	10	18.9	
Difficulties to refrain from smoking							
No	98	73.1	112	70.4	38	71.7	0.878
Yes	36	26.9	47	29.6	15	28.3	
Longest period of abstinence							
Days	44	23.4	59	29.7	22	35.5	0.189
Months	96	51.1	85	42.7	22	35.5	
Years	48	25.5	55	27.6	18	29.0	
Ever tried to quit (current smokers only)							
Yes	91	65.4	94	56.4	36	69.2	0.139
No	42	31.6	67	41.6	16	30.8	

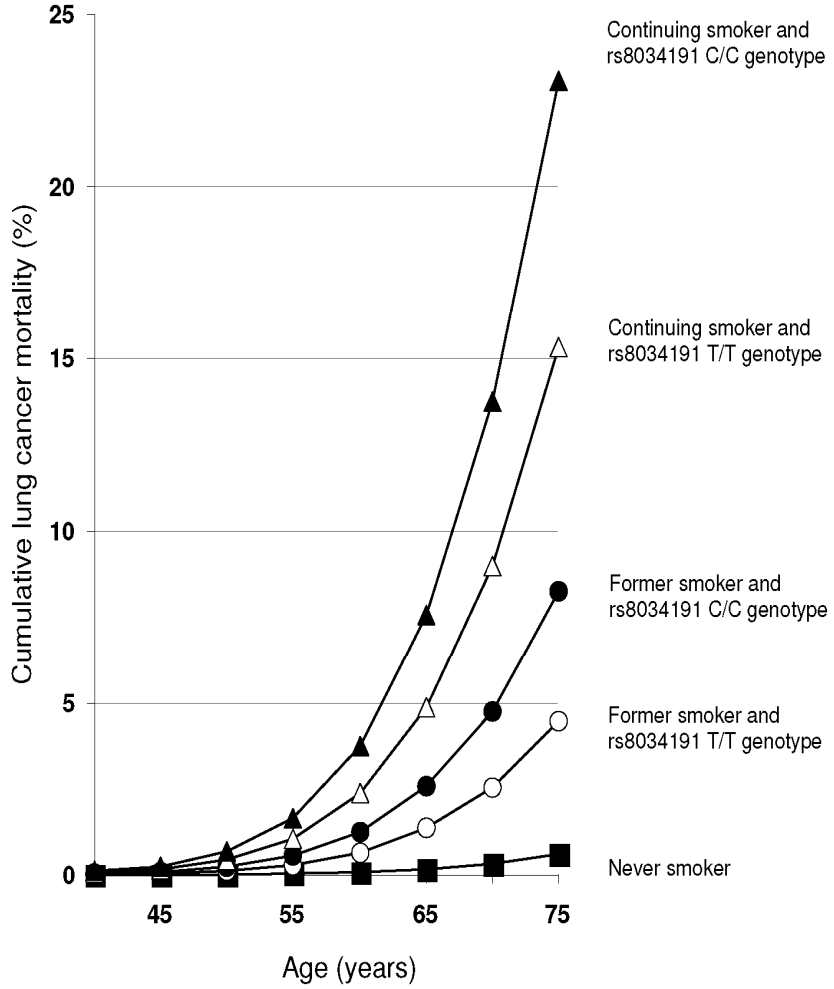
no association with head and neck cancer in 2,262 cases vs. 2,137 controls

International Agency for Research on Cancer

Centre International de Recherche sur le Cancer



Impact on life-time risk of lung cancer death: Example from Poland

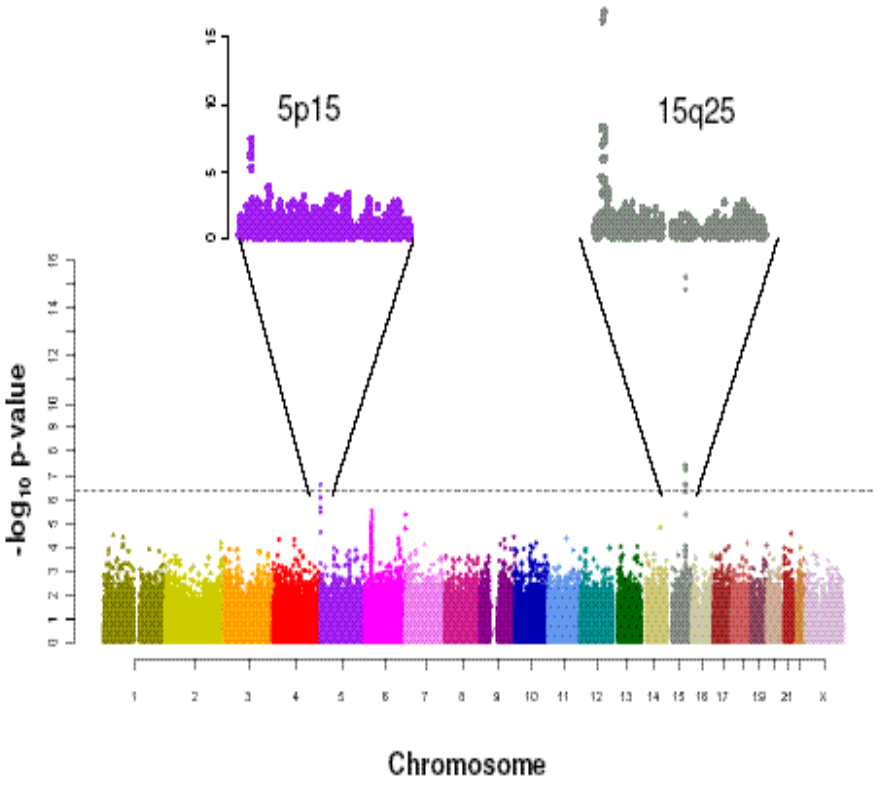


CC genotype occurs in about 12% of population

Centre International de Recherche sur le Cancer

Lung cancer 2nd phase GWA scan

McKay etal. Nat Gen 03 Nov 2008

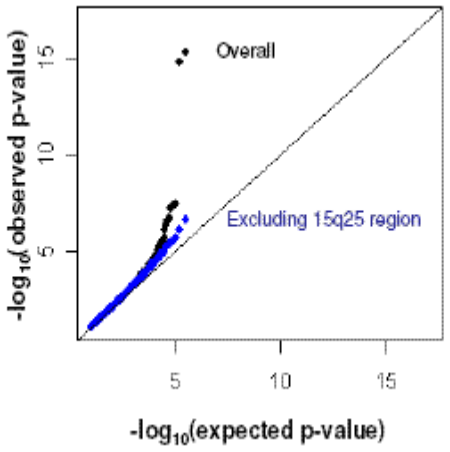


3259 Caucasian lung cancer cases

5573 matched controls

310,023 SNPs

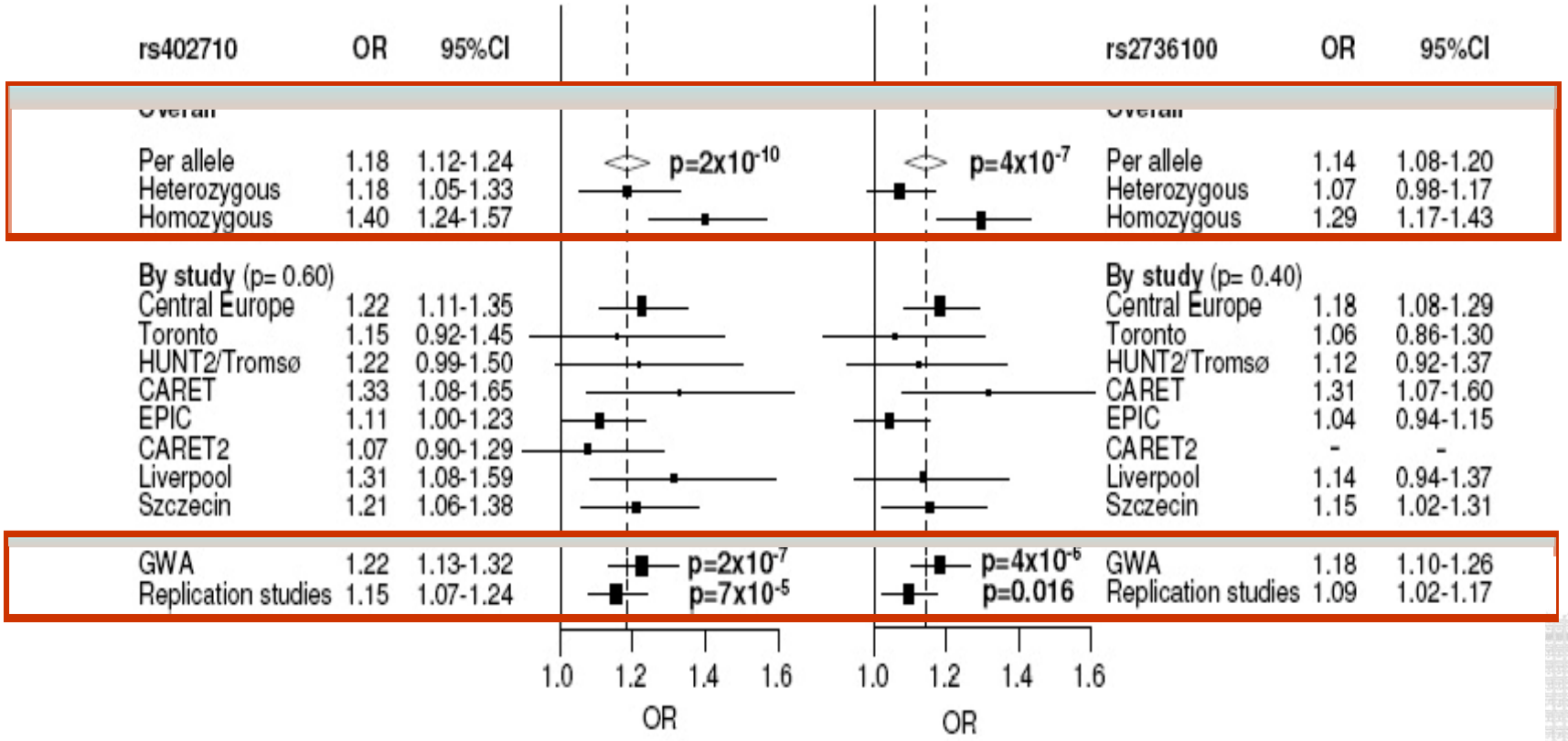
Genome-wide significance $p<5x10^{-7}$



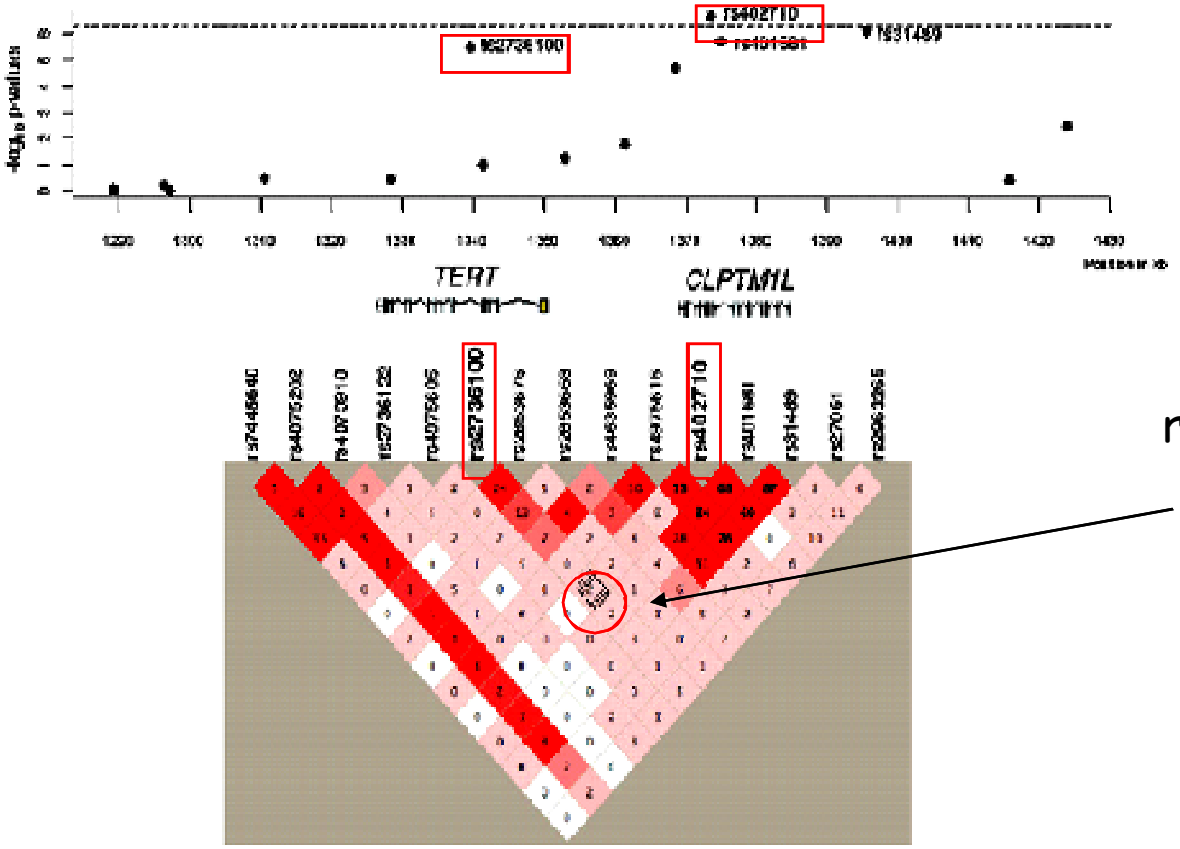
International Agency for Research on Cancer

Centre International de Recherche sur le Cancer

Detection & replication of lung cancer association on chromosome 5p



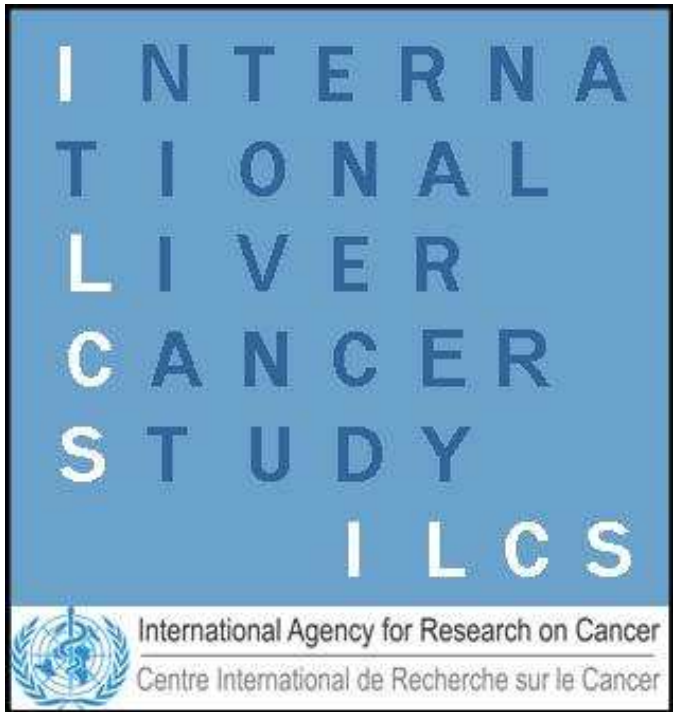
Chromosome 5p15.33 association in lung cancer



rs2736100-
rs402710
D'=0.231,
r2=0.026

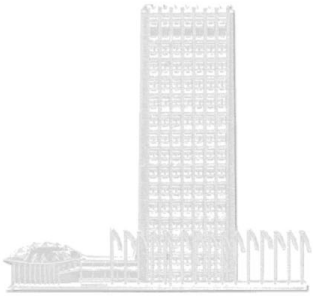
TERT = reverse transcriptase component of telomerase
CLPTM1L = cleft lip and palate transmembrane 1 like

International Liver Cancer Study

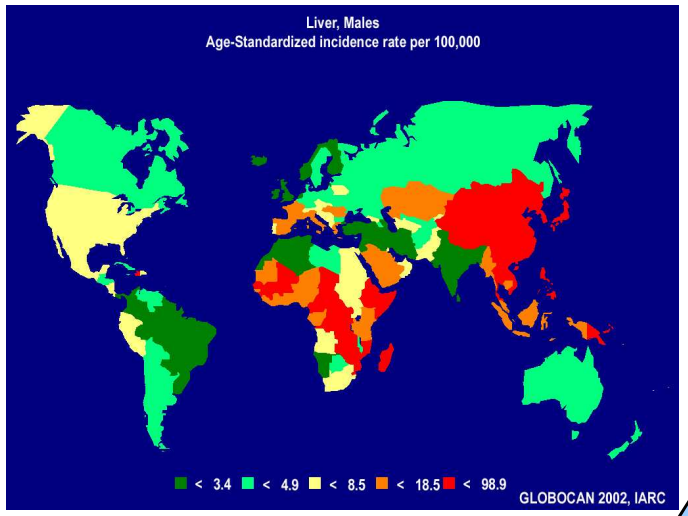


http://ilcs.iarc.fr

International Agency for Research on Cancer
Centre International de Recherche sur le Cancer



Integrated approach for biomarker discovery: liver carcinogenesis



Genomics

Molecular pathology

Plasma Proteomics

Epigenetics in plasma and tumor

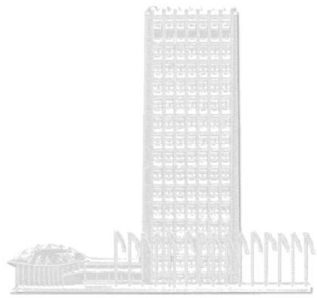
HBV/HCV sequencing

Mutations in Plasma DNA

Definition

- International Liver Cancer Study (ILCS): a global network of case/case, case-control and cohort studies to assess the geographic pathological, virological and molecular diversity of HCC in order to (1) discover biomarkers for prevention and early detection; (2) develop approaches for reducinhg the global mortality by HCC

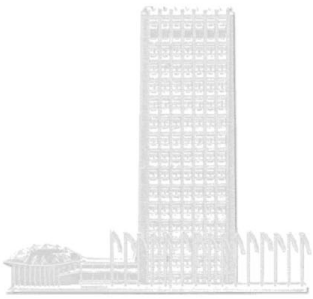
- *ILCS is an initiative National Cancer Centers:* Kam Man Hui (Singapore), Tadao Kakizoe (Japan), David Kerr (United Kingdom), Thiravud Khuhaprema (Thailand), Srivatanakul Petcharin (Thailand), Khee Chee Soo (Singapore), Robert Thomas (Australia)



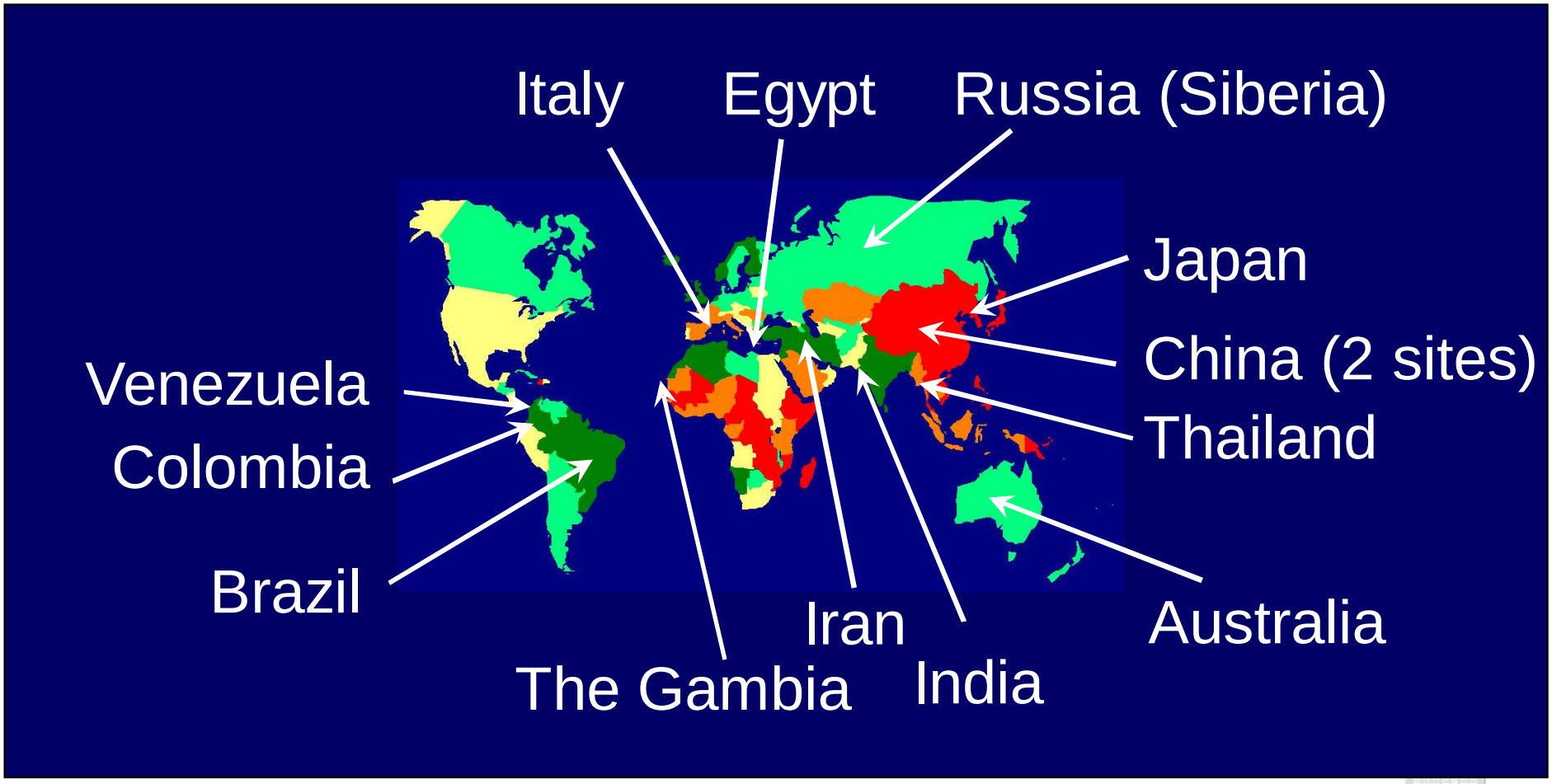
Structure of the Progamme

Three-phases programme

1. Standardized protocol for constitution of an annotated biological resource
2. Protocols for large-scale, international studies for biomarker discovery
3. Validation studies in various epidemiological and clinical contexts

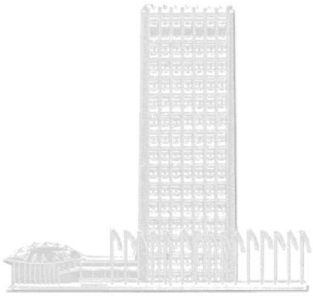


ILCS: Global outreach



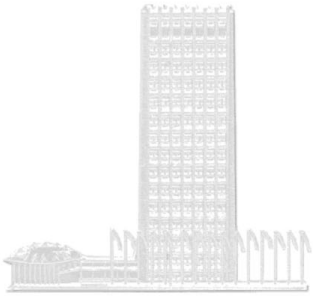
Data and Specimen Collection

- Hospital setup
- Based on simple case definition (Clinical diagnosis, ultrasonography, AFP)
- Common clinical and epidemiological questionnaire
- Case-only design
- Case-control design
- Cohort study design
- Blood (fractionated and frozen OR Guthrie cards)
- Biospies (FFPE, frozen, RNA later)
- Surgical specimen with adjacent tissue (Frozen)



Biobankig approach

- Each participating center develops a collection which is commensurate with its logistical and infrastructure capacity
- The basic, standard level is blood collection on Guthrie Card, which has been validated for DNA extraction, HBsAg, AFP and HBV genotyping
- For tissues, the basic standards are RNA-later and FFPE biopsies
- Only specialized centers collect samples for cryopreservation

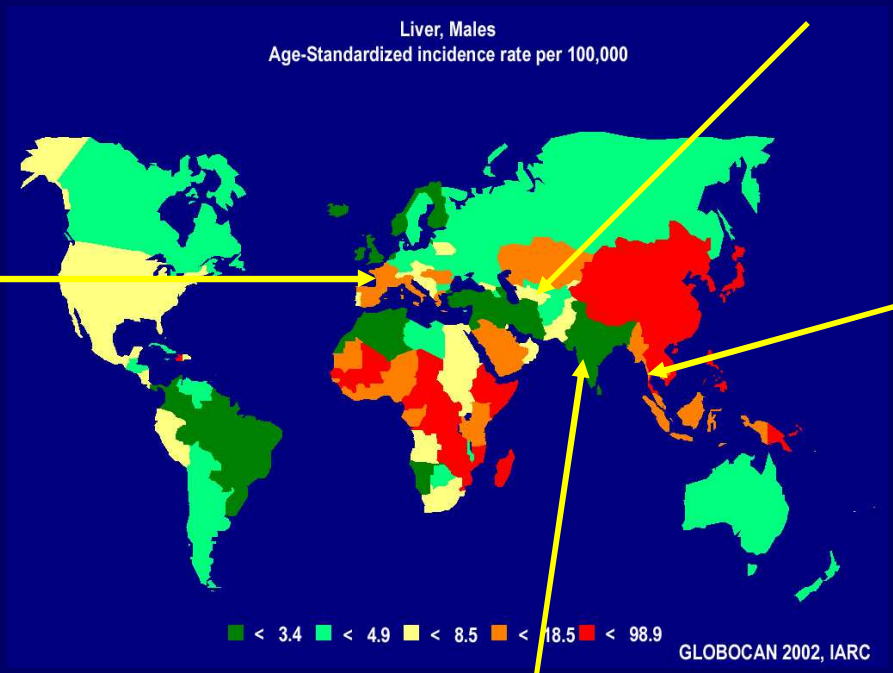


Cohort Studies

Northern Iran: 50 000 subjects, 10 years follow-up, initiated 2001; Blood, Biospies; active follow-up

EPIC cohort: 111 cases of HCC, prospective design, Blood, follow-up by cancer registration, Started in 1992

Italy (Naples)

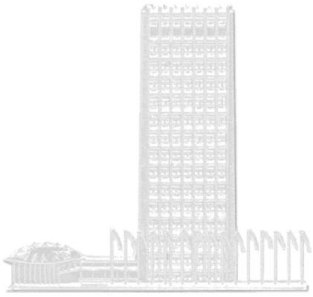


Thailand cohort: 1810 males aged +35, HBV carriers, recruited between 1987 and 1992, followed up since. blood samples; clinical follow-up; blood, frozen surgical samples

India (Trivandrum) cohort: 600 males and females, HBV carriers, recruited since 2003. New cohort of up to 2000 carriers to be recruited as nested study in a large population based cohort study; Blood, Tissues, active follow-up

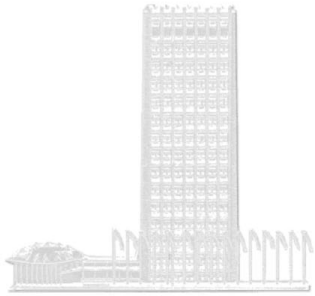
Proteomics Pilot

- Aim: Discovery and validation of plasma proteomic biomarkers for early detection in particular in a non-cirrhosis context
- Within current HPPP coordinated by Dr Laura Beretta (Fred Hutchinson Cancer Research Center, Seattle)
- Studies to date are focusing on HCC in US, mostly in relation with HCV and cirrhosis
- Within ILCS, we are contributing additional collections from orther parts of the world: The Gambia (done); Thailand (ongoing); Colombia (planned); Egypt (under discussion)



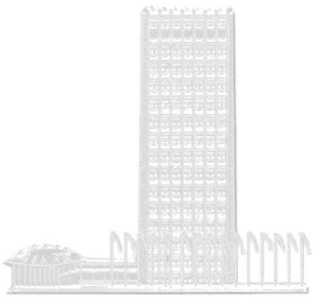
Proteomics Pilot: Validation phase

- Jan-June 2009: data mining, prioritization of markers of interest
- Constitution of validation series: HCV/Cirrhosis (USA); The Gambia (Case-control); Europe (EPIC, prospective); Thailand (HBV carriers, prospective); India (HBV carriers, prospective)
- First step of biomarker assessment and extension to other ILCS series

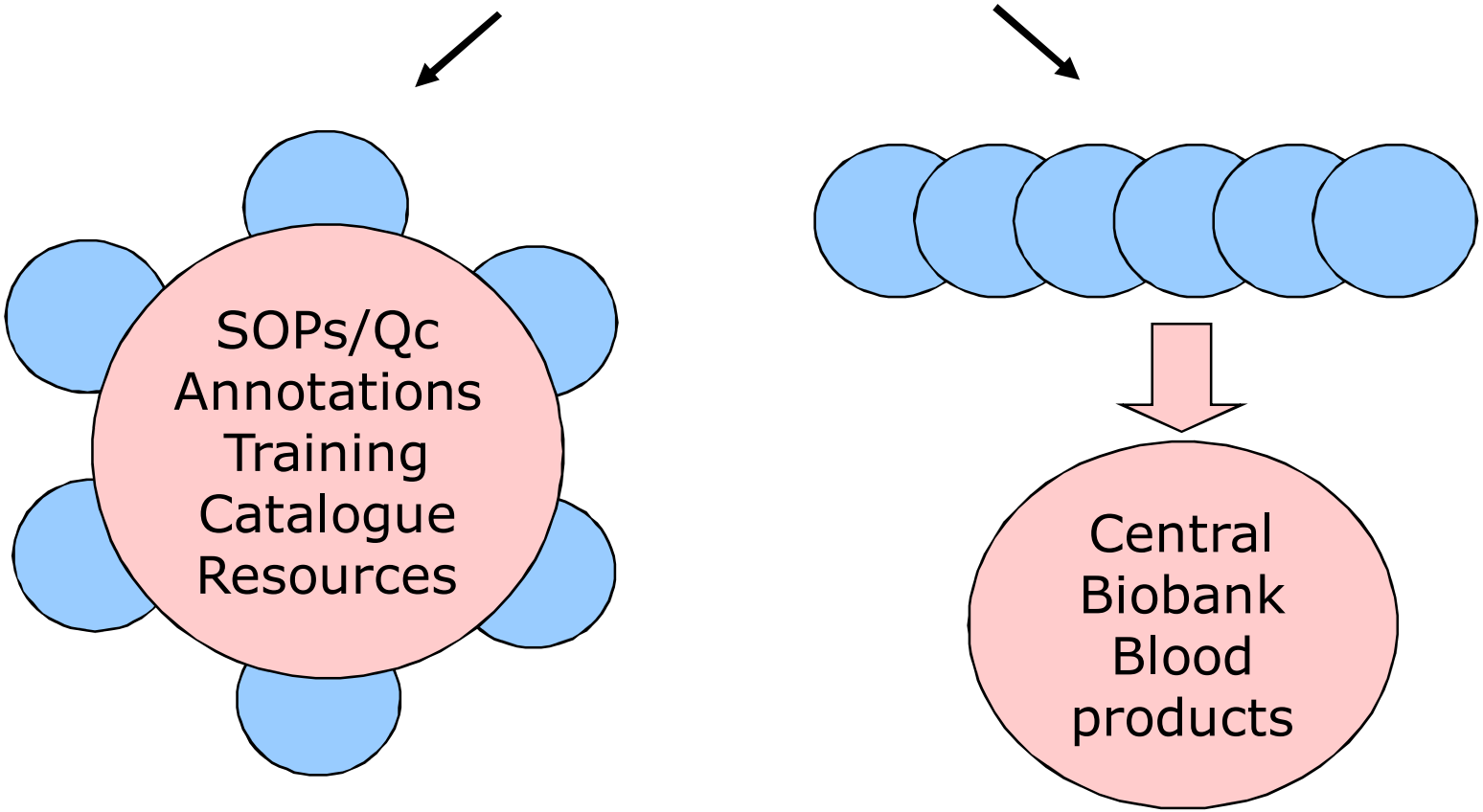


An Integrated Model

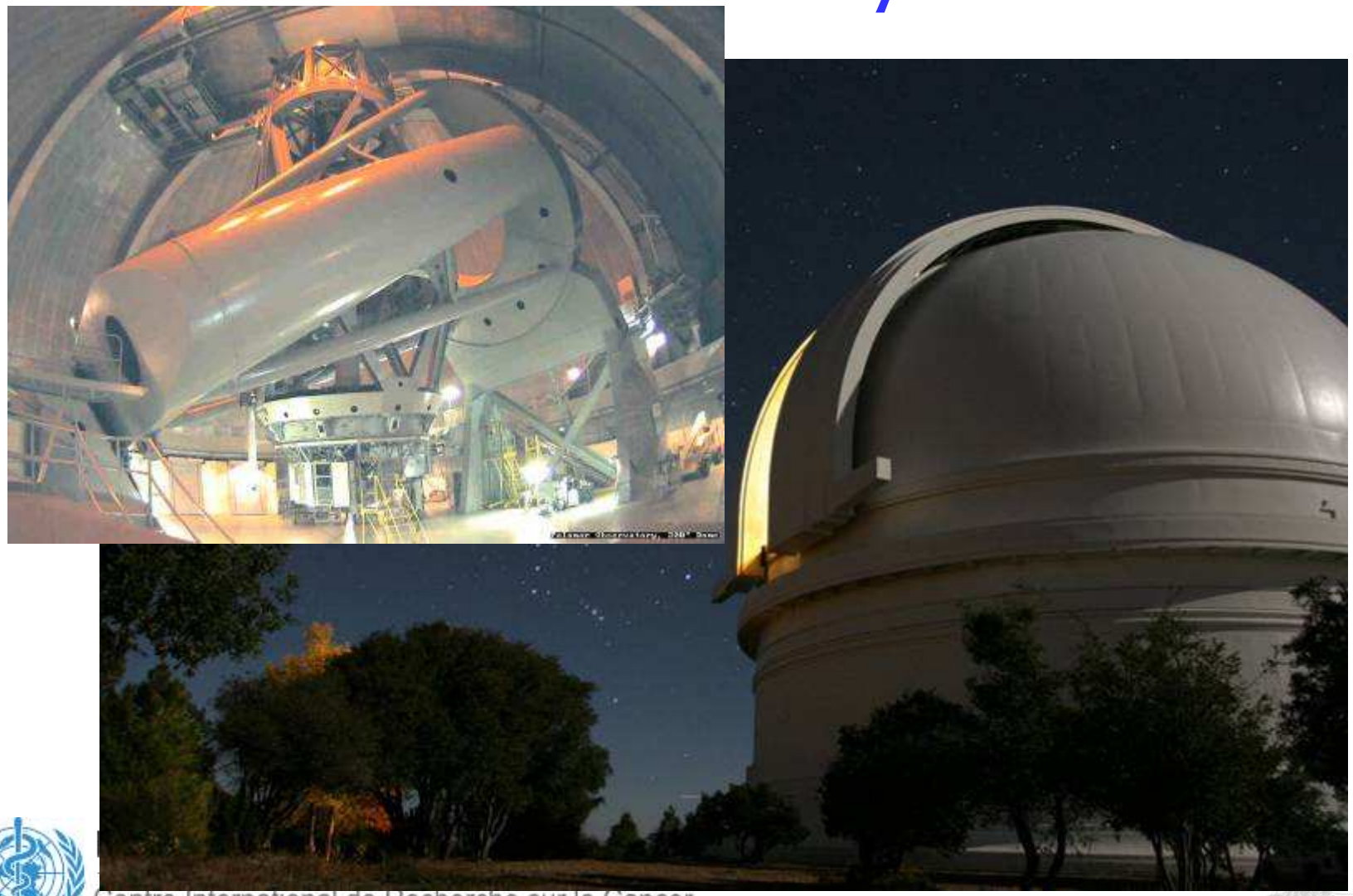
International Agency for Research on Cancer
Centre International de Recherche sur le Cancer



Tissues/blood



Building Large Instruments for Discovery



Further information

Biobanks at IARC

Elodie Caboux, BRC Manager, BRC@iarc.fr
Carole Cravotto, BBMRI and Biobank coordination
Dominique Bouchard, BRC secretary and administration
BRC staff (7 technicians)

Scientific supervision: Pierre Hainaut
Scientific strategy: Markus Pasterk

To receive an copy of these slides:
hainaut@iarc.fr

To access IARC technical guidelines:

http://www.iarc.fr/IARCPress/pdfs/StandardsBRC/index.php