193 Yurdaydin C, Bozkaya H, Cetinkaya H *et al.* Lamivudine vs lamivudine and interferon combination treatment of HBeAg(-) chronic hepatitis B. *J Viral Hepat* 2005; **12**: 262–8.
194 Sarin SK, Kumar M, Kumar R *et al.* Higher efficacy of sequential therapy with interferon-alpha and lamivudine combination compared to lamivudine monotherapy in HBeAg positive chronic hepatitis B patients. *Am J Gastroenterol* 2005; **100**: 2463–71.
195 Yang YF, Zhao W, Zhong YD, Xia HM, Shen L, Zhang N. Interferon therapy in chronic hepatitis B reduces progression to cirrhosis and hepatocellular carcinoma: a meta-analysis. *J Viral Hepat* 2009; **16**: 265–2671.
196 Lin SM, Sheen IS, Chien RN *et al.* Long-term beneficial effect of interferon therapy in patients with chronic hepatitis B virus infection. *Hepatology* 1999; **29**: 971–5.
197 van Zonneveld M, Honkoop P, Hansen BE *et al.* Long-term follow-up of alpha-interferon treatment of patients with chronic hepatitis B. *Hepatology* 2004; **39**: 804–10.

Hepatology Research

Hepatology Research 2011; **41**: 936–945 doi: 10.1111/j.1872-034X.2011.00859.x

Original Article

Easy-to-use phylogenetic analysis system for hepatitis B virus infection

Masaya Sugiyama,[1,2]* Ayano Inui,[3] Tadasu Shin-I,[1] Haruki Komatsu,[3] Motokazu Mukaide,[1,4] Naohiko Masaki,[1] Kazumoto Murata,[1] Kiyoaki Ito,[1] Makoto Nakanishi,[2] Tomoo Fujisawa[3] and Masashi Mizokami[1]

[1]*The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa,* [2]*Department of Biochemistry and Cell Biology, Nagoya City University Graduate School of Medical Sciences, Nagoya,* [3]*Department of Pediatrics, Eastern Yokohama Hospital, Yokohama, and* [4]*SRL, Inc. Tokyo, Japan*

Aim: The molecular phylogenetic analysis has been broadly applied to clinical and virological study. However, the appropriate settings and application of calculation parameters are difficult for non-specialists of molecular genetics. In the present study, the phylogenetic analysis tool was developed for the easy determination of genotypes and transmission route.

Methods: A total of 23 patients of 10 families infected with hepatitis B virus (HBV) were enrolled and expected to undergo intrafamilial transmission. The extracted HBV DNA were amplified and sequenced in a region of the S gene.

Results: The software to automatically classify query sequence was constructed and installed on the Hepatitis Virus Database (HVDB). Reference sequences were retrieved from HVDB, which contained major genotypes from A to H. Multiple-alignments using CLUSTAL W were performed before the genetic distance matrix was calculated with the six-parameter method. The phylogenetic tree was output by the neighbor-joining method. User interface using WWW-browser was also developed for intuitive control. This system was named as the easy-to-use phylogenetic analysis system (E-PAS). Twenty-three sera of 10 families were analyzed to evaluate E-PAS. The queries obtained from nine families were genotype C and were located in one cluster per family. However, one patient of a family was classified into the cluster different from her family, suggesting that E-PAS detected the sample distinct from that of her family on the transmission route.

Conclusions: The E-PAS to output phylogenetic tree was developed since requisite material was sequence data only. E-PAS could expand to determine HBV genotypes as well as transmission routes.

Key words: database, genotype, hepatitis B virus, intrafamilial transmission, phylogenetic analysis

INTRODUCTION

HEPATITIS B VIRUS (HBV) infects approximately 350 million people worldwide. Chronic HBV infection causes liver cirrhosis and liver cancer. In Japan, chronic hepatitis B patients are estimated to be approximately one million.[1] As HBV has high infectivity, almost all advanced countries have launched the universal infant immunization program against HBV. The patients with HB antigen seropositive, fulminant hepatitis and HCC have substantially declined in these countries.[2,3] The current immunization strategy represents favorable effects to prevent the HBV transmission. In countries without the universal vaccination program, however, substantial intrafamilial transmission and horizontal transmission have been reported.[4] In Japan, only high-risk infants born to chronic HBV-infected mothers have been given the HBV vaccine according to a selective vaccination policy of Japanese governments since 1986. This strategy has led to successful reduction of HBV carrier infants[5] since the major route of transmission has been perinatal transmission,[6,7] and horizontal transmission in early childhood has occurred as a result of close family contact.[8–10] However, the main

Correspondence: Dr Masashi Mizokami, The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, 1-7-1, Kohnodai, Ichikawa 272-8516, Japan. Email: mmizokami@hospk.ncgm.go.jp
*JSPS Research Fellow.
Received 14 April 2011; revision 13 June 2011; accepted 15 June 2011.

route of acute hepatitis B in adult is sexual transmission in countries without a universal vaccination program.

Hepatitis B virus genotypes are identified worldwide and classified into at least eight genotypes (A–H) on the basis of a divergence of 8% or more of the entire nucleotide sequences.[1,11–13] In a Japanese population of chronic hepatitis B, the distribution of major HBV genotypes (HBV/A, B, C, and D) was reported to be 1.7%, 12.2%, 84.7%, and 0.4% respectively.[14] However, the prevalence of HBV/A increased to approximately 40% in acute HBV infection, and 4.3% in chronic HBV infection in Japan.[15] The main transmission routes of HBV/A are a contact among men who have sex with men as well as among members of a heterosexual population.[16–19] The expansion of HBV infection has been led by the sexual transmission of HBV/A since little infection by contaminated medical materials has recently been reported in Japan.

As previously reported, phylogenetic analyses based on virus genome revealed the transmission route.[20] The identification of the transmission route provides beneficial information for epidemiologic study and healthcare reform. Although phylogenetic analysis is gradually known in this field, the handlings of the data are difficult to achieve for researchers who are not familiar with genetic analysis. In this study, we have developed a novel computing system to output phylogenetic trees automatically when users simply input their sequence data. The result from this system also shows the type of transmission route such as intrafamilial (mother-to-child, father-to-child, or child-to-child) or horizontal in the general population. For this purpose, sera collected from 10 families with possible intrafamilial transmission were used to reveal their accurate transmission route.

METHODS

Patients

TEN FAMILIES CONSISTING of 23 patients being inactive carriers or having chronic hepatitis B were enrolled in this study after 32 individuals of 10 families

Table 1 Characteristics of intrafamilial transmission cases

Family	Feature	Age	Sex	HBV DNA (Log copies/mL)	Status 1	Status 2	ID
FM1	C	11	F	>9	eAg +	IC	1
	Mo	40	F	8.7	eAg +	IC	2
FM2	C	7	M	>9	eAg +	IC	3
	Mo/S1	41	F	8.6	eAg +	CH	4
	S2	27	F	<2.6	eAb +	IC	5
	S3	30	F	3	eAb +	IC	6
FM3	C	3	F	>9	eAg +	IC	7
	Mo	37	F	>9	eAg +	IC	8
FM4	C1	15	M	>9	eAg +	IC	9
	C2	20	F	>9	eAg +	IC	10
	Mo	47	F	>9	eAg +	IC	11
FM5	C	15	M	5.9	eAb +	IC	12
	Fa	46	M	4.2	eAb +	IC	13
FM6	C	12	M	>9	eAg +	CH	14
	Mo	36	F	>9	eAg +	CH	15
FM7	C	2	F	2.8	eAb +	IC	16
	Mo	33	F	>9	eAg +	IC	17
FM8	C	2	M	>9	eAg +	IC	18
	Mo	41	F	>9	eAg +	IC	19
FM9	C	7	M	3.5	eAb +	IC	20
	Mo	39	F	>9	eAg +	CH	21
FM10	C1	4	M	>9	eAg +	CH	22
	C2	2	F	<2.6	eAb +	IC	23

All patients were hepatitis B surface antigen (HBsAg) positive. Feature: C, children; Mo, mother; Fa, father; S, sister. Sex: M, male; F, female. Status1: Status of HB antigen or antibody except for HbsAg. Status2: IC, inactive carrier; CH, chronic hapatitis. ID 4, 5, 6 are sister and ID 4 is mother of ID 3. ID 22 and 23 are brother and sister.

were tested on HB markers (Table 1). All patients were hepatitis B surface antigen (HBsAg) positive, and the family cases had a possibility of intrafamilial transmission. The histories of familial clustering of HBV infection and hepatitis B vaccination were recorded. The serum samples obtained from all patients were tested for hepatitis B surface antigen (HBsAg), hepatitis B surface antibody (anti-HBs), and hepatitis B core antibody (anti-HBc). Four sera of chronic hepatitis B patients were collected for operation-checks of the developed system before a trial of 23 samples. Serum samples were divided into aliquots and kept at −80°C until testing. The study protocol conformed to the 1975 declaration of Helsinki and was approved by the ethics committees of the respective institutions. Every patient or his/her next of kin gave informed consent to the purpose of the study. Consent of children for participating in the study was filled by their parents.

Serological testing

Their sera were tested for alanine aminotransferase (ALT), and hepatitis B e antigen (HBeAg) and hepatitis B s antigen (HBsAg), as well as antibodies to HBeAg (anti-HBe) and HBsAg (anti-HBs) (Dinabot, Tokyo, Japan). Antibodies to HBcAg (anti-HBc) were tested by ARCHITECT (Abbott Japan, Tokyo, Japan). The inactive carrier state was defined by the presence of HBV surface antigen (HBsAg) with normal ALT levels over 1 year (examined at least four times at 3-month intervals). Chronic hepatitis was defined by elevated ALT levels (>1.5 times the upper limit of normal [35 IU/L]) persisting over 6 months (with at least three bimonthly tests). HBV DNA levels in sera were quantitated with a commercial kit (Taqman Real-time polymerase chain reaction [PCR] or Amplicor HBV Monitor; Roche Diagnostics, Basel, Switzerland) with a detection range from 2.6 to 9 log copies/mL.

Viral DNA extraction

HBV DNA was extracted from 200 μL of serum using QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA, USA) according to manufacturer's instruction. The extracted DNA was used for amplification and direct sequencing of S gene as described below.

HBV DNA sequencing

The target of S gene (255 bp, nucleotide positions 458–712) was amplified by nested PCR and sequenced to detect in high sensitivity. The forward primers of S gene were HBS/F2, 5′- AGGTATGTTGCCCGTTTGTC -3′ for the outer set and HBS/F1, 5′- GTATGTTGCCCGTTT GTCCT -3′ for the inner set. The reverse primers of the gene were HBS/R2, 5′- AAAGCCCTACGAACCACTGA –3′ for the outer set and HBS/R1, 5′- AAGCCCTACG AACCACTGAA-3′ for the inner set. Nested PCRs were performed with these primers for 35 cycles (95°C, 15 s; 58°C, 30 s; 72°C, 30 s) in 1st and 2nd PCR using Veriti (Applied Biosystems, Foster City, CA, USA). The PCR products were sequenced on both strands with the BigDye Terminator V3.1 cycle sequencing kit (Applied Biosystems) with the same primers used for the 2nd PCR. The sequencing products were analyzed with an ABI 3130xl DNA analyzer (Applied Biosystems). The obtained sequences were aligned with GenBank sequences corresponding to HBV genotypes.

Database and the system for phylogenetic analysis

The determination of transmission routes using phylogenetic analyses is considered based on age difference, mutation rate, clinical background, and the amplicon size between queries. The determination of non-intrafamilial transmission between queries was performed using a significance level of $P < 10^{-99}$. The criterion was calculated by the simulation of random sampling from Hepatitis Virus Database (HVDB) and the genetic factors described above. The presently developed system, the easy-to-use phylogenetic analysis system (E-PAS) for the determination of genotypes and/or transmission route is implemented in the account mode of the HVDB (http://s2as02.genes.nig.ac.jp). We recommend that researchers contact the web master before use.

Table 2 Reference data for the analysis of transmission root

Genotype	Country	Number of references
A	Foreign	2
B	Japan	24
C	Japan	64
D	Japan	1
E	Foreign	1
F	Foreign	1
G	Foreign	1
H	Japan	3

a

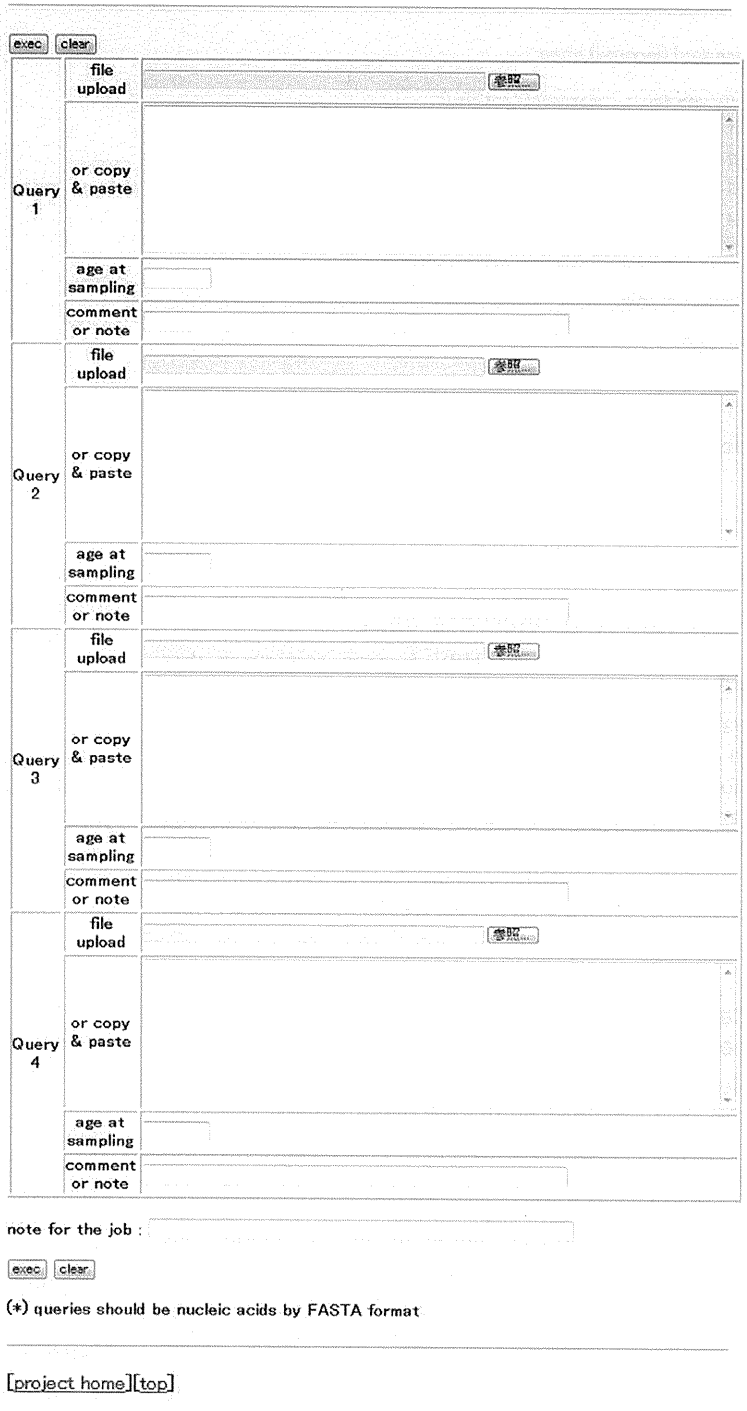


Figure 1 WWW-based user interface. To avoid bias such as age and selective pressure of antiviral treatment, adjustment parameters are prepared to obtain appropriate results. (a) The input field of query sequences and patient's age when serum was collected is shown. (b) The upper field configures masked sites for the elimination of drug-resistant sites. The lower field is a mutation rate of hepatitis B virus (HBV).

b

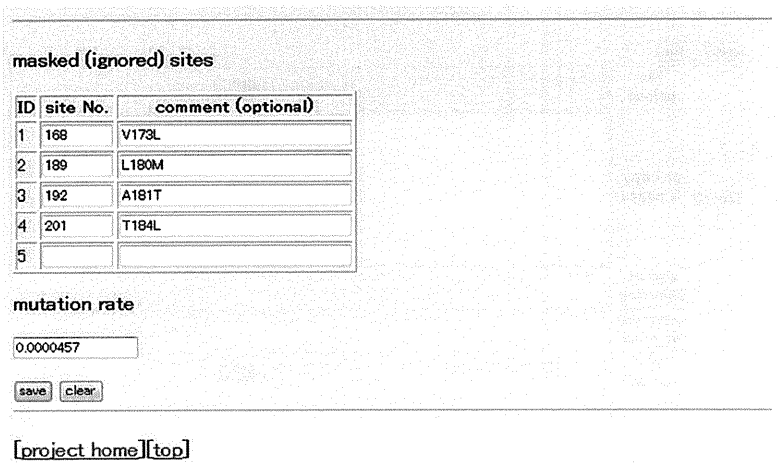


Figure 1 *Continued.*

RESULTS

Preparation of reference data

A DATASET OF all HBV S gene sequences that contain sequence variations in Japanese patients was retrieved from the HVDB[21] for reference sequences to classify suspicious samples of intrafamilial transmission. A representative sequence was selected among the same sequences in a cluster to optimize phylogenetic analysis. Full annotations of all sequences were also extracted from HVDB. The data of major genotypes were registered into the system. The results of the dataset were shown in Table 2.

Calculation for the classification of new sequences

New sequences input in query columns were calculated and then classified by the following process (Fig. 1a). The query sequences were added to the reference database described above, and the overall sequences were multiple-aligned by CLUSTAL W[22]. More precise alignment than direct alignment among queries was obtained regardless of the sequence quality of queries. Then, the genetic distance matrix was calculated from the alignment with the six-parameter method,[23] using only the sites shared by all the sequences. Finally the phylogenetic tree was constructed from the matrix using the neighbor-joining method,[24] which showed the relation among the queries, between query and published entries. The above calculation was conducted without using nucleotide positions of drug-induced resistant mutations to avoid selective pressure of antiviral treatment (Fig. 1b). As shown in Figure 1b, the mutation site of Lamivudin, Adefovir diproxil, and Entecavir resistance were masked in the present study along with the previous study.[25,26] We have estimated the substitution rate for HBV to be 4.57×10^{-5} per site per year.[27] In addition, the adjustment of age factor was calculated in consideration of genetic distance between queries for precise phylogenetic analysis (Fig. 1a). The distribution of nucleotide differences could be generally approximated by Poisson distribution and its average difference site number in the PCR amplicon of this study was 4.67 bases. The probability that two independent samples had the same sequence by chance was estimated to be less than 1%, and then the false positive data were seldom output in the intrafamilial transmission case. This system showed a sensitivity of 100% by calculating independent data retrieved from Osiowy *et al.*[28] The paper provided sequence data of a 25-year period obtained from eight asymptomatic carriers of the HBV genotype B in 1979 and 2004. For the calculation of the specificity, random sampling was performed using HVDB data. The specificity of this system was 98.6% by sampling 72 data.

Users interface

We also developed a WWW-based user interface to conduct the analysis easily. Researchers get two types of results in one analysis after they upload two or more query sequences through WWW browsers. One is a report of nucleotide difference number between

queries. This shows an indication of whether queries are similar enough or not. The other is a phylogenetic tree that shows the relation among queries and reference data.

Determination of transmission route in family cases

Four independent samples of chronic hepatitis B obtained from our hospital were tested before the analysis of suspicious intrafamilial transmission cases. Each of these four sequences was classified into distinct cluster (data not shown). As the E-PAS worked well, the analysis of suspicious intrafamilial transmission cases was examined to reveal whether two (or more) sequences obtained from a family were from intrafamilial transmission or not. The analysis was carried out by the process described above. The queries in the multiple alignments were compared between queries to count different nucleotide sites. Then, the comparison is done using the high-quality regions of the queries by ignoring ambiguous base ("N" base) sites. If three or more queries of a family were entered in the system, the comparison was calculated for each pair of queries. Table 1 shows characteristics of patients

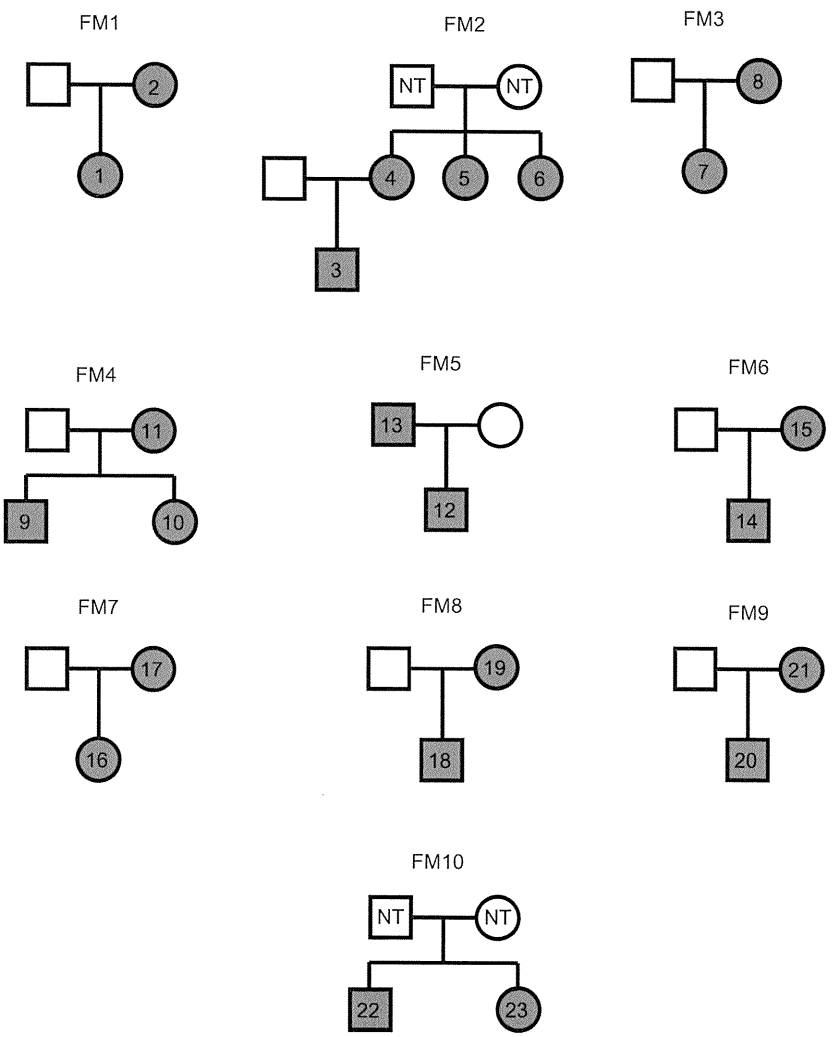
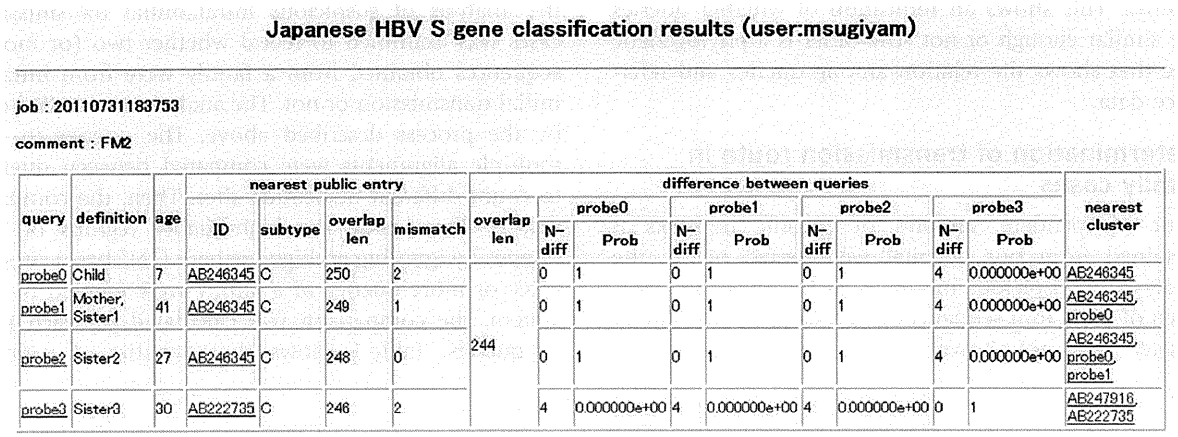


Figure 2 Family trees for 10 families with clustering hepatitis B virus (HBV) infection. Gray color indicates HBsAg positive, and white color indicates hepatitis B surface antigen (HBsAg) negative. Patients without HB antigen information are described as not tested (NT). Squares indicates male sex, circles indicates female sex.



Japanese HBV S gene classification results (user:msugiyam)

job : 20110731183753

comment : FM2

query	definition	age	nearest public entry				difference between queries								nearest cluster
			ID	subtype	overlap len	mismatch	overlap len	probe0		probe1		probe2		probe3	
								N-diff	Prob	N-diff	Prob	N-diff	Prob	N-diff	Prob
probe0	Child	7	AB246345	C	250	2	244	0	1	0	1	0	1	4	0.000000e+00	AB246345
probe1	Mother, Sister1	41	AB246345	C	249	1		0	1	0	1	0	1	4	0.000000e+00	AB246345, probe0
probe2	Sister2	27	AB246345	C	248	0		0	1	0	1	0	1	4	0.000000e+00	AB246345, probe0, probe1
probe3	Sister3	30	AB222735	C	246	2		4	0.000000e+00	4	0.000000e+00	4	0.000000e+00	0	1	AB247916, AB222735

Figure 3 WWW-based user interface and the results of the matrix calculation by the software. The queries registered on the system were multi-aligned and analyzed for the calculation of mismatched nucleotides between the members of the family. This figure shows a representative result using FM2 data. The definition column presents sample information. The age column shows the point when their sera was collected. The nearest public entry shows the data locating near the query and displays the genotype, the length of query, and mismatched nucleotides compared with the nearest reference entry. The column of the difference between queries represents the mismatched nucleotides between the queries. The N-diff column shows the mismatched nucleotides between the queries. The probability value of the difference between two queries displays the prob column.

infected with HBV. These 10 families were analyzed to determine the transmission route using the present system. Predictably, two or three sequence data obtained from one family were classified into one cluster except for FM2 case (data not shown), and then the transmission routes of these nine families were intrafamilial as was expected.

A case of multiple transmission routes in a family

Interestingly, the result of FM2 case was different from the others. As shown in Table 1 and Figure 2, the family consisted of four carriers (three sisters and one of their sons). The number of nucleotide differences was calculated after the multiple-alignment by CLUSTAL W. Sample ID 6 (sister 3, S3) showed four mismatched nucleotides compared with the other members in the family (see the column of difference between queries in Fig. 3), and then the phylogenetic tree was described using E-PAS in HVDB. As shown in Figure 4, the sequence of ID 6 was classified into the cluster far from the other member, suggesting that ID 6 was infected with a different HBV clone from the other member. The probability value (abbreviated as Prob in Fig. 3) of intrafamilial transmission between two queries is displayed in the browser.

DISCUSSION

WE HAVE DEVELOPED an easy-to-use system for the calculation of phylogenetic analysis. The present system (E-PAS) was constructed for non-specialist users of genetics as well as specialists and would provide the appropriate results with several adjusted parameters by just preparing sequence data. Therefore, this system is recommended to users who want to easily investigate transmission routes of samples. In the present study, we have applied the E-PAS to the determination of transmission routes in families. Almost all of the family-cases showed intrafamilial transmission because the queries belonging to one family were classified in same cluster or a neighboring position. In the analysis between families, these queries certainly were located in the distinct cluster. One patient (sample ID 6) showed quite different sequences from the other members of her family. We did not determine whether her transmission route was vertical or horizontal as data of her parents on HBV infection were not provided. If both parents were HBV carriers, the different sequences could be detectable among ID 4, 5, and 6 as an intrafamilial transmission from father or mother to child, or she was infected by a vertical transmission from non-familial member.

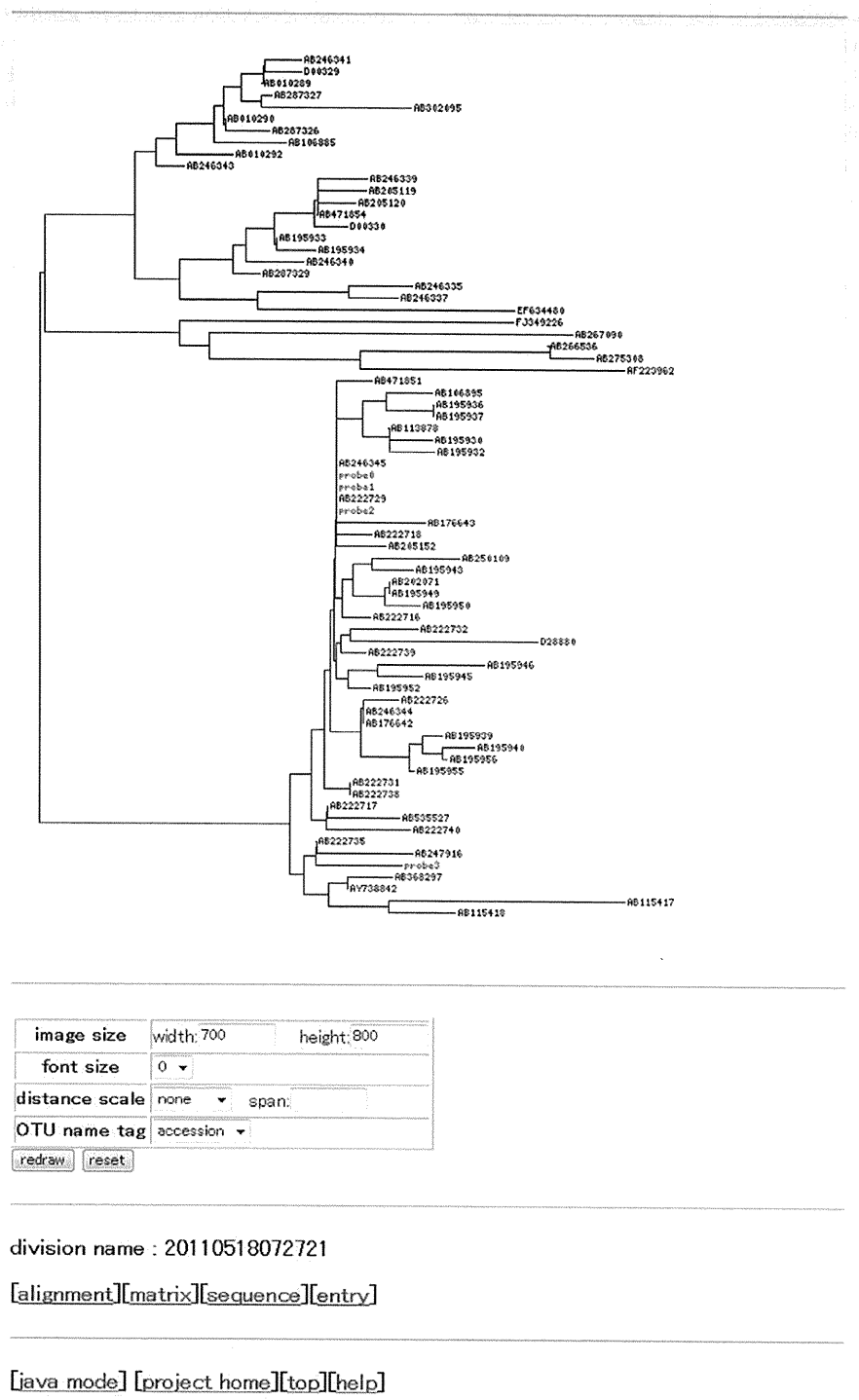


Figure 4 WWW-based user interface and the results of the phylogenetic tree. The phylogenetic tree was calculated based on the matrix data of mismatched nucleotides. A representative result using FM2 data was shown. The character in red indicates the queries. Three members of FM2 had the same sequence in the S region and were classified into the same cluster, whereas sample ID 6 (showed as probe 3 in tree) was classified into another cluster.

The primer set designed in the S region revealed high sensitive amplification because samples with low titer of HBV DNA under the limit of measurement were amplified effectively and the single band was detected in almost all samples by only the 1st PCR. Although we have estimated the substitution rate for HBV to be 4.57×10^{-5} per site per year,[27] the amplified region is a relatively highly conserved area even under antiviral drug pressure[29,30] and the region shows the genotype-dependent sequence. Therefore, a transmission route of sample ID 6 could be divided from that of the other family members. In line with this result, the determination of transmission routes and/or genotyping was achieved sufficiently even if HBV DNA was negative in conventional diagnosis. Therefore, these procedures should be determined with great caution to avoid contamination by trained researchers. The software based on a WWW-browser was customized to fit a Japanese population for quick calculation by reducing entry sequences of foreign origin. Additional datasets are under development because the present system focused on genotype B and C that are originally prevalent in Japan.

This is a versatile tool as the route of suspicious intrafamilial transmission and the genotype were determined in this assay. HBV/A has recently spread out among the young heterosexual population in addition to men who have sex with men.[16,18,19] The determination of transmission routes could contribute to epidemiological research and/or the decision of government policy to prevent the spread of HBV/A. Generally, acute hepatitis B in adulthood becomes chronic in only <1% of cases, which is much less frequent than that in Europe and the United States (approximately 10%). Since HBV/A has been more frequently observed recently in Japan, especially in metropolitan areas, HBV/A infection could provide an increased risk of chronic diseases. Further, HBV/A might be a majority of HBV infection in Japan. Reference sequence of HBV/A registered on this system could be also the appropriate reference for genotyping of HBV/A extracted from Japanese because the HBV/A2 isolates detected in Japan were homologous to those from Europe and the United States in the phylogenetic analysis.[19]

In conclusion, we have developed the easy-to-use system, E-PAS, to determine the transmission route. E-PAS separated one case infected with a clone that was different from the family, which had been expected to be intrafamilial transmission. The system is expected to accelerate the phylogenetic analysis for researchers or physicians who are not familiar with molecular genetics. Although the E-PAS has been primarily developed for Japanese patients, we are planning to extend the application for other patients by collecting foreign samples and reference sequences in addition to improve the calculation method of genetic distance.

ACKNOWLEDGMENTS

THIS WORK WAS supported by a Grant-in-Aid from the Ministry of Health Labor and Welfare of Japan and a Grant-in-Aid for JSPS Fellows of the Ministry of Education, Culture, Sports, Science, and Technology of Japan. The authors thank Professor Takashi Gojobori, National Institute of Genetics and Professor Yoshiyuki Suzuki, Nagoya City University for their professional opinion and technical support. We also thank Ms. Masako Awano, Ms. Miwa Noda, Ms. Ryuko Izumida, Ms. Mikako Kajio, and Ms. Mika Saito for their secretarial support. Laboratory work performed by Ms. Sachiko Sato, Ms. Miki Yoshida, Ms. Chieko Haga, Ms. Mami Ohashi, and Ms. Naomi Nomura is highly appreciated.

REFERENCES

1 Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *J Gen Virol* 2002; **83**: 2059–73.
2 Chang MH. Impact of hepatitis B vaccination on hepatitis B disease and nucleic acid testing in high-prevalence populations. *J Clin Virol* 2006; **36** (Suppl 1): S45–50.
3 Chen DS. Hepatitis B vaccination: the key towards elimination and eradication of hepatitis B. *J Hepatol* 2009; **50**: 805–16.
4 Kao JH, Chen DS. Global control of hepatitis B virus infection. *Lancet Infect Dis* 2002; **2**: 395–403.
5 Tada H, Uga N, Fuse Y *et al.* Prevention of perinatal transmission of hepatitis B virus carrier state. *Acta Paediatr Jpn* 1992; **34**: 656–9.
6 Stevens CE, Beasley RP, Tsui J, Lee WC. Vertical transmission of hepatitis B antigen in Taiwan. *N Engl J Med* 1975; **292**: 771–4.
7 Sung JL, Chen DS. Maternal transmission of hepatitis B surface antigen in patients with hepatocellular carcinoma in Taiwan. *Scand J Gastroenterol* 1980; **15**: 321–4.
8 Szmuness W, Prince AM, Hirsch RL, Brotman B. Familial clustering of hepatitis B infection. *N Engl J Med* 1973; **289**: 1162–6.
9 Lok AS, Lai CL, Wu PC, Wong VC, Yeoh EK, Lin HJ. Hepatitis B virus infection in Chinese families in Hong Kong. *Am J Epidemiol* 1987; **126**: 492–9.

10 Dumpis U, Holmes EC, Mendy M *et al.* Transmission of hepatitis B virus infection in Gambian families revealed by phylogenetic analysis. *J Hepatol* 2001; **35**: 99–104.
11 Okamoto H, Tsuda F, Sakugawa H *et al.* Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. *J Gen Virol* 1988; **69** (Pt 10): 2575–83.
12 Norder H, Courouce AM, Magnius LO. Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* 1994; **198**: 489–503.
13 Stuyver L, De Gendt S, Van Geyt C *et al.* A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. *J Gen Virol* 2000; **81**: 67–74.
14 Orito E, Ichida T, Sakugawa H *et al.* Geographic distribution of hepatitis B virus (HBV) genotype in patients with chronic HBV infection in Japan. *Hepatology* 2001; **34**: 590–4.
15 Kobayashi M, Ikeda K, Arase Y *et al.* Change of hepatitis B virus genotypes in acute and chronic infections in Japan. *J Med Virol* 2008; **80**: 1880–4.
16 Yano K, Tamada Y, Yatsuhashi H *et al.* Dynamic epidemiology of acute viral hepatitis in Japan. *Intervirology* 2010; **53**: 70–5.
17 Suzuki Y, Kobayashi M, Ikeda K *et al.* Persistence of acute infection with hepatitis B virus genotype A and treatment in Japan. *J Med Virol* 2005; **76**: 33–9.
18 Yotsuyanagi H, Okuse C, Yasuda K *et al.* Distinct geographic distributions of hepatitis B virus genotypes in patients with acute infection in Japan. *J Med Virol* 2005; **77**: 39–46.
19 Matsuura K, Tanaka Y, Hige S *et al.* Distribution of hepatitis B virus genotypes among patients with chronic infection in Japan shifting toward an increase of genotype A. *J Clin Microbiol* 2009; **47**: 1476–83.
20 Mizokami M, Gojobori T, Lau JY. Molecular evolutionary virology: its application to hepatitis C virus. *Gastroenterology* 1994; **107**: 1181–2.
21 Shin IT, Tanaka Y, Tateno Y, Mizokami M. Development and public release of a comprehensive hepatitis virus database. *Hepatol Res* 2008; **38**: 234–43.
22 Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; **22**: 4673–80.
23 Gojobori T, Ishii K, Nei M. Estimation of average number of nucleotide substitutions when the rate of substitution varies with nucleotide. *J Mol Evol* 1982; **18**: 414–23.
24 Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; **4**: 406–25.
25 Locarnini SA, Yuen L. Molecular genesis of drug-resistant and vaccine-escape HBV mutants. *Antivir Ther* 2010; **15**: 451–61.
26 Zoulim F, Locarnini S. Hepatitis B virus resistance to nucleos(t)ide analogues. *Gastroenterology* 2009; **137**: 1593–608 e1-2.
27 Orito E, Mizokami M, Ina Y *et al.* Host-independent evolution and a genetic classification of the hepadnavirus family based on nucleotide sequences. *Proc Natl Acad Sci USA* 1989; **86**: 7059–62.
28 Osiowy C, Giles E, Tanaka Y, Mizokami M, Minuk GY. Molecular evolution of hepatitis B virus over 25 years. *J Virol* 2006; **80**: 10307–14.
29 Mizokami M, Orito E, Ohba K, Ikeo K, Lau JY, Gojobori T. Constrained evolution with respect to gene overlap of hepatitis B virus. *J Mol Evol* 1997; **44** (Suppl 1): S83–90.
30 Pallier C, Castera L, Soulier A *et al.* Dynamics of hepatitis B virus resistance to lamivudine. *J Virol* 2006; **80**: 643–53.

Bousquet *et al. Genome Medicine* 2011, **3**:43
http://genomemedicine.com/content/3/7/43

Genome Medicine

CORRESPONDENCE Open Access

Systems medicine and integrated care to combat chronic noncommunicable diseases

Jean Bousquet[1*], Josep M Anto[2], Peter J Sterk[3], Ian M Adcock[4], Kian Fan Chung[5], Josep Roca[6], Alvar Agusti[6], Chris Brightling[7], Anne Cambon-Thomsen[8], Alfredo Cesario[9], Sonia Abdelhak[10], Stylianos E Antonarakis[11], Antoine Avignon[12], Andrea Ballabio[13], Eugenio Baraldi[14], Alexander Baranov[15], Thomas Bieber[16], Joël Bockaert[17], Samir Brahmachari[18], Christian Brambilla[19], Jacques Bringer[20], Michel Dauzat[21], Ingemar Ernberg[22], Leonardo Fabbri[23], Philippe Froguel[24], David Galas[25], Takashi Gojobori[26], Peter Hunter[27], Christian Jorgensen[28], Francine Kauffmann[29], Philippe Kourilsky[30], Marek L Kowalski[31], Doron Lancet[32], Claude Le Pen[33], Jacques Mallet[34], Bongani Mayosi[35], Jacques Mercier[36], Andres Metspalu[37], Joseph H Nadeau[25], Grégory Ninot[38], Denis Noble[39], Mehmet Öztürk[40], Susanna Palkonen[41], Christian Préfaut[36], Klaus Rabe[42], Eric Renard[20], Richard G Roberts[43], Boleslav Samolinski[44], Holger J Schünemann[45], Hans-Uwe Simon[46], Marcelo Bento Soares[47], Giulio Superti-Furga[48], Jesper Tegner[49], Sergio Verjovski-Almeida[50], Peter Wellstead[51], Olaf Wolkenhauer[52], Emiel Wouters[53], Rudi Balling[54], Anthony J Brookes[55], Dominique Charron[56], Christophe Pison[57,58], Zhu Chen[59], Leroy Hood[25] and Charles Auffray[56,57,58,60,61]

Abstract

We propose an innovative, integrated, cost-effective health system to combat major non-communicable diseases (NCDs), including cardiovascular, chronic respiratory, metabolic, rheumatologic and neurologic disorders and cancers, which together are the predominant health problem of the 21st century. This proposed holistic strategy involves comprehensive patient-centered integrated care and multi-scale, multi-modal and multi-level systems approaches to tackle NCDs as a common group of diseases. Rather than studying each disease individually, it will take into account their intertwined gene-environment, socio-economic interactions and co-morbidities that lead to individual-specific complex phenotypes. It will implement a road map for predictive, preventive, personalized and participatory (P4) medicine based on a robust and extensive knowledge management infrastructure that contains individual patient information. It will be supported by strategic partnerships involving all stakeholders, including general practitioners associated with patient-centered care. This systems medicine strategy, which will take a holistic approach to disease, is designed to allow the results to be used globally, taking into account the needs and specificities of local economies and health systems.

*Correspondence: jean.bousquet@inserm.fr
[1]Department of Respiratory Diseases, Arnaud de Villeneuve Hospital, CHU Montpellier, INSERM CESP U1018, Villejuif, France
Full list of author information is available at the end of the article

Non-communicable diseases, the major global health problem of the century

Chronic diseases are disorders of long duration and generally slow progression [1]. They include four major non-communicable diseases (NCDs) listed by the World Health Organization (WHO) [2] – cardiovascular diseases, cancer, chronic respiratory diseases and diabetes – as well as other NCDs, such as neuropsychiatric disorders [3] and arthritis. As survival rates have improved for infectious and genetic diseases, chronic diseases have come to include communicable diseases (such as HIV/AIDS) and genetic disorders (such as cystic fibrosis). NCDs represent the major global health problem of the 21st century [4,5]; they affect all age groups [6] and their burden is greater than that of infectious diseases. NCDs are the world leading cause of disease burden and mortality [2] and are increasing in prevalence and burden [7], even in low- and middle-income countries [8]. Costs incurred by uncontrolled NCDs are substantial, especially in underserved populations [9] and low- and middle-income countries [10,11]. NCDs are an under-appreciated cause of poverty and hinder economic development [11]. Importantly, management of NCDs has recently been prioritized globally (Box 1).

Chronic diseases are caused by complex gene-environment interactions acting across the lifespan from the fetus to old age (Figure 1). In this context, 'environment' includes risk and protective factors associated with environment and lifestyle, such as

Box 1: Priorities for the prevention and control of NCDs

May 2008: 61st World Health Assembly. WHO recommended a worldwide priority policy on NCD prevention and control (2008 to 2013), including cardiovascular disease, cancer, chronic respiratory diseases [101] and diabetes, not least because they often have common environmental risk factors [2].

May 2010: United Nations (UN) General Assembly unanimously adopted Resolution A/RES/64/265: 'Tackling NCDs constitutes one of the major challenges for sustainable development in the 21st century' [102].

December 2010: the Council of the European Union adopted conclusions based on innovative and global approaches for NCDs in public health and healthcare systems to further develop population-based and patient-centered policies [1].

2010: US Center for Disease Control and Prevention (CDC) [103] says that 'an essential strategy for keeping older adults healthy is preventing NCDs and reducing associated complications'.

19 September 2011: UN General Assembly symposium on NCDs.

tobacco, nutrition, indoor and outdoor air pollution and sedentary life [2].

Socio-economic determinants are intertwined with the onset, progression, severity and control of NCDs. There are functional interdependencies between molecular components, reflecting complex network perturbations that link cells, tissues and organs [12]. Early life events are crucial in the generation of NCDs, and aging increases disease complexity, adding, for example, tissue and cell senescence [13]. Comorbidity refers to the co-existence of two or more diseases or conditions in the same individual that have similar risk factors and/or mechanisms. Most people with NCDs suffer from two or more diseases [14]. Co-morbidity and multi-morbidity are common signatures of NCDs and are associated with worse health outcomes [15], complex pharmacological interventions and clinical management, and increased healthcare costs [16]. However, little is known about how NCDs truly cluster at the genetic, molecular or mechanistic levels, and there is scant understanding of

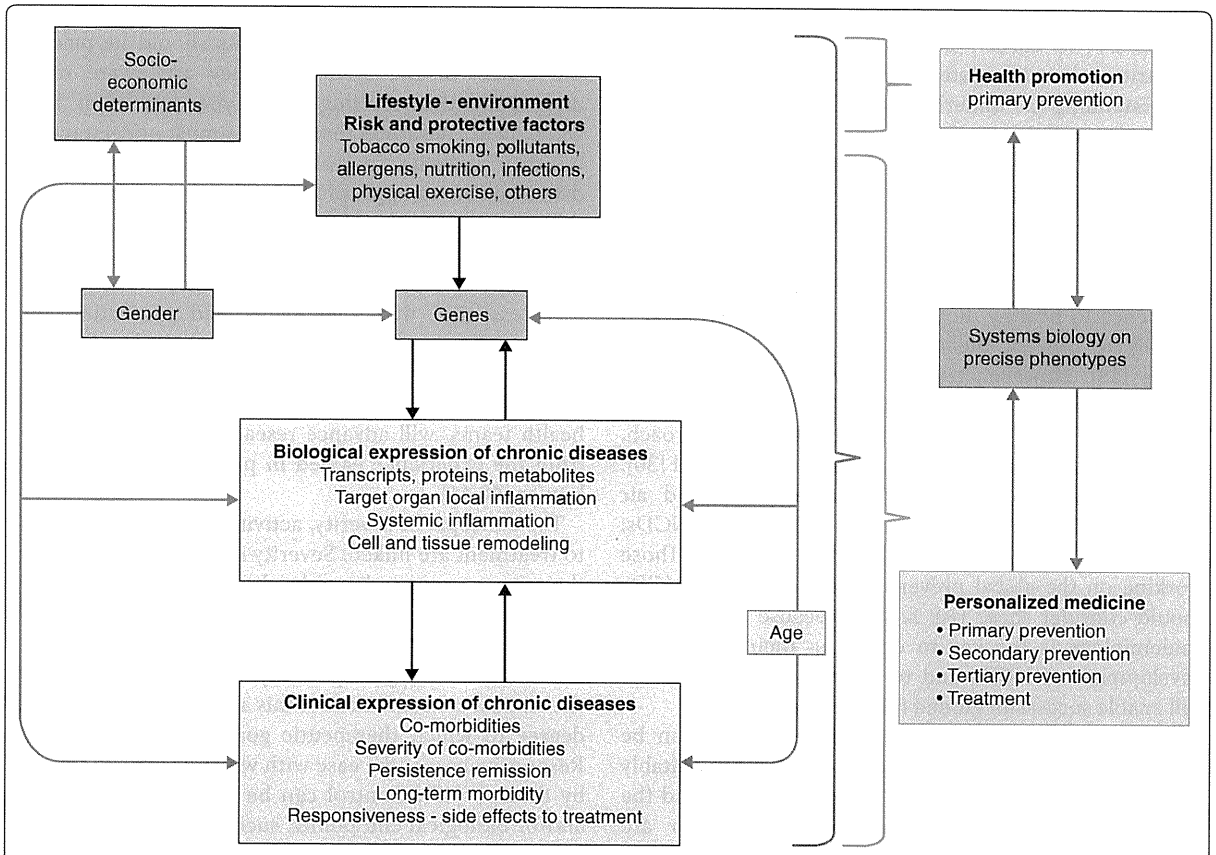


Figure 1. NCDs are associated with complex gene-environment interactions modulated by socio-economic determinants, psychological factors, age and gender. The products of these interactions lead to the biological expression of NCDs and further to their clinical expression with co-morbidities. A new definition of NCD phenotypes is needed to understand how a network of molecular and environmental factors can lead to complex clinical outcomes of NCDs for prevention and control.

how specific combinations of NCDs influence prognosis and treatment [16].

NCDs are multi-factorial. In addition to environmental factors and increased life expectancy, intrinsic host responses, such as local and systemic inflammation, immune responses and remodeling [17], have key roles in the initiation and persistence of diseases and co-morbidities. The recent increase in NCDs has been associated in part with biodiversity loss [18], socio-economic inequities linked with climate change, and loss of natural environments [19]. A more comprehensive understanding of these links will make it possible to propose more effective primary prevention strategies. The *in utero* environment is an important determinant of adult NCDs, including diabetes [20], coronary heart diseases [21], and asthma [22] or chronic obstructive pulmonary disease (COPD) [23]. Mechanistic links have been proposed that involve fetal expression of genes that are conserved across species, epigenetic mechanisms [22,24], early and maternal life infections, and/or environmental exposures. These need to be understood better [25], as early interventions may have the potential to reduce NCD burden [26].

Nutrition is a key determinant of health and NCDs. Understanding the underlying complexities of metabolic responses and pathophysiology is needed. Loss of biodiversity in food organisms causes micronutrient and vitamin deficiencies, and obesity and related chronic and degenerative diseases are a formidable challenge [27]. Nutritional intervention in early childhood may help prevent autoimmune diseases [28], and adoption and adherence to healthy diet recommendations are needed globally to prevent the onset and facilitate control of NCDs [29]. However, trying to change lifestyles using public health efforts remains a major challenge, and an interdisciplinary social and behavioral approach, including the cultural aspects of nutrition, is needed [30]. Tobacco use [31], biomass fuel combustion and air pollution [32] are among the major risk factors for NCDs; these act as early as *in utero* and in early life. Those working on the global prevention and control of NCDs should consider these risk factors because translational epidemiology is the key to exploring their role in the development of NCDs and to devising approaches that will enable successful guided interventions [33].

The development of a society, rich or poor, can be judged by the health of its population, how equitably health is distributed across the social spectrum, and the degree of protection provided to people who are disadvantaged by illness. Effective action against NCDs needs to include understanding of the social and economic determinants and their modification (Figure 1) [34]. Indeed, best-practice interventions targeted at coronary risk factors eliminate most socioeconomic differences that affect coronary heart disease mortality, and this should serve as an example to follow for other NCDs [35]. In May 2009, the 62nd WHO Assembly recommended re-orienting health systems globally to promote primary healthcare as the most cost-effective strategy [36]. Healthcare often focuses on single diseases, advanced technology, biomedical interventions and specialist care. Most healthcare takes place in primary care settings [37], with emphasis on providing a complete range of care, from home to hospital, and on investing resources rationally. Fragmenting care can reduce the ability of primary care clinicians to ensure that patient care is comprehensive, integrated, holistic, and coordinated [38], and to decide whether a person has a significant disease or temporary symptoms [39].

A proposal for multidisciplinary patient-centered management of chronic NCDs

We recommend that, to determine measures of disease severity and control, effective interventions and studies should be built around carefully phenotyped patients (Figure 2) and strictly follow carefully crafted methodological standards. Patients should be placed at the center of the system; if they are aware of and understand the resulting phenotype data, their health will benefit. We stress that patients must understand that it is their societal responsibility to make their anonymized data available to appropriate scientists and physicians so that the latter can create the predictive medicine of the future that will transform the health of their children and grandchildren. For patients to adopt this approach, it is essential that laws be passed protecting them against abuse of their personal data by insurance companies, health authorities or employers. This approach to patient-centeredness, if aided by community health teams, will advance research. It may also benefit from the experience gained in patient-centered medical homes [40,41].

The concepts of severity, activity, control and response to treatment are linked. Severity is the loss of function in the target organs induced by disease and may vary over time; as it may also vary with age, this needs to be regularly re-evaluated. Activity is the current level of activation of biological pathways causing the disease and the clinical consequences of this activation. Control is the degree to which therapeutic goals are being met [42]. Responsiveness is the ease with which control is achieved by therapy [43]. Control can be achieved using clinical and/or biological end points, such as glycemic control in diabetes [44]. Careful monitoring of co-factors, such as compliance, and of unavoidable risk factors is needed. The uniform definition of severe asthma presented to WHO is based on this approach [45] and therefore provides a model to assess NCD severity (Figure 3).

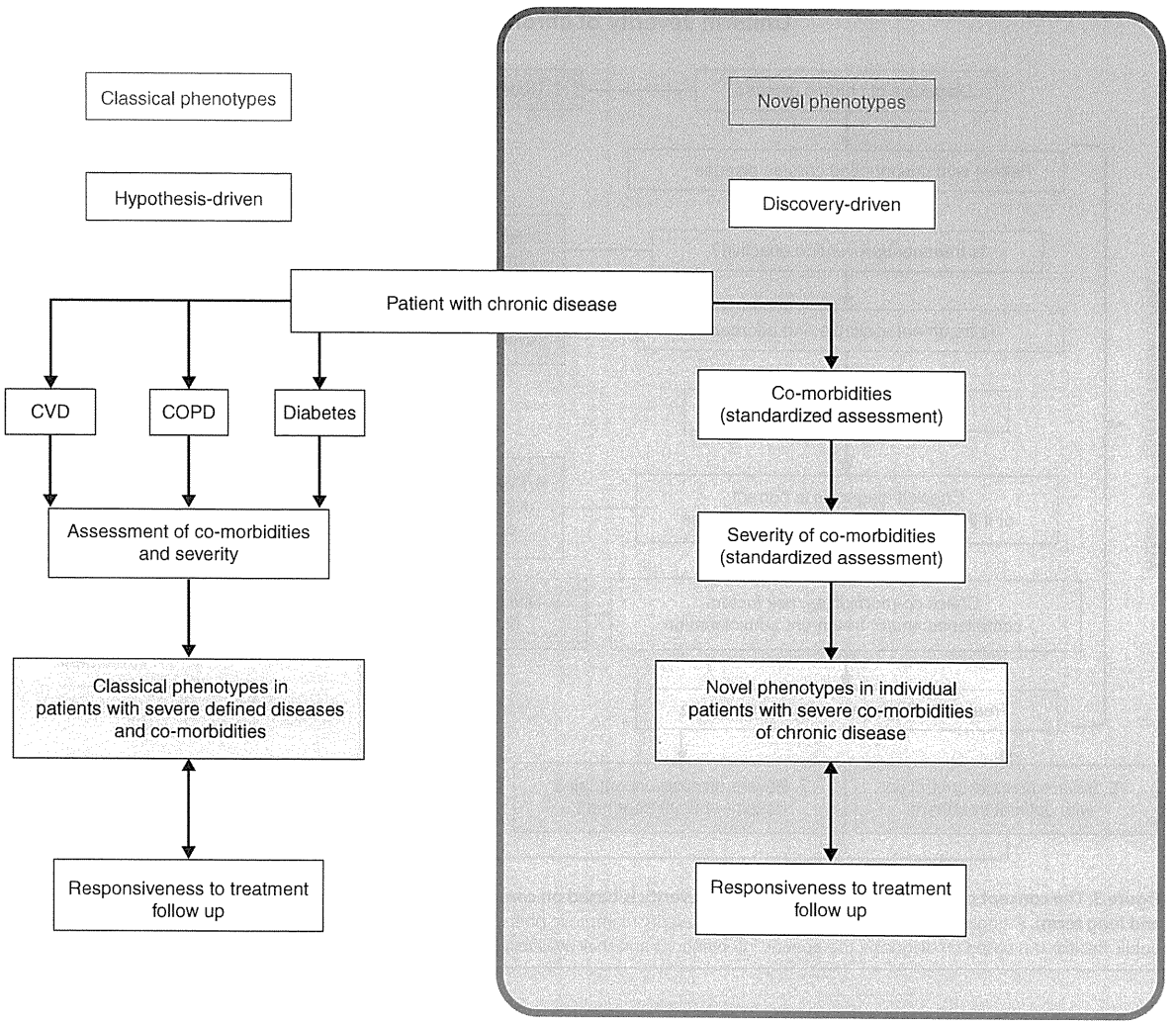


Figure 2. Classical phenotypes are based on *a priori* ontologies (cardiovascular disease, chronic obstructive pulmonary disease (COPD) and type 2 diabetes), and new phenotypes are based on statistical modeling of all the complex components of NCD onset, persistence and prognosis.

Information and communication technologies (ICT) are needed for the implementation of integrated care in a systems medicine approach to enable prospective follow-up of the patients. Home telemonitoring is promising [46] and should be explored further because continuous and precise monitoring makes each individual clinical history a valuable source of comprehensive information. More user-friendly and efficient ICT platforms are needed that include shared decision making, the process by which a healthcare choice is made jointly by the practitioner and the patient [47]. Ideally, an innovative patient management program would combine ICT, shared decision making and personalized education of the patient (and caregiver) about multidisciplinary approaches. The content, acceptance and effectiveness of such approaches should be tested to ensure that the autonomy, quality of life and capacity of patients are respected and enhanced, and that their values and preferences dominate decision making [48]. Practice-based inter-professional collaborations is also key to improving healthcare processes and outcomes [49]. Qualitative assessment will provide insight into how interventions affect collaboration and how improved collaboration contributes to changes in outcomes.

Thus, we propose that NCD management should move towards holistic multi-modal integrated care, and

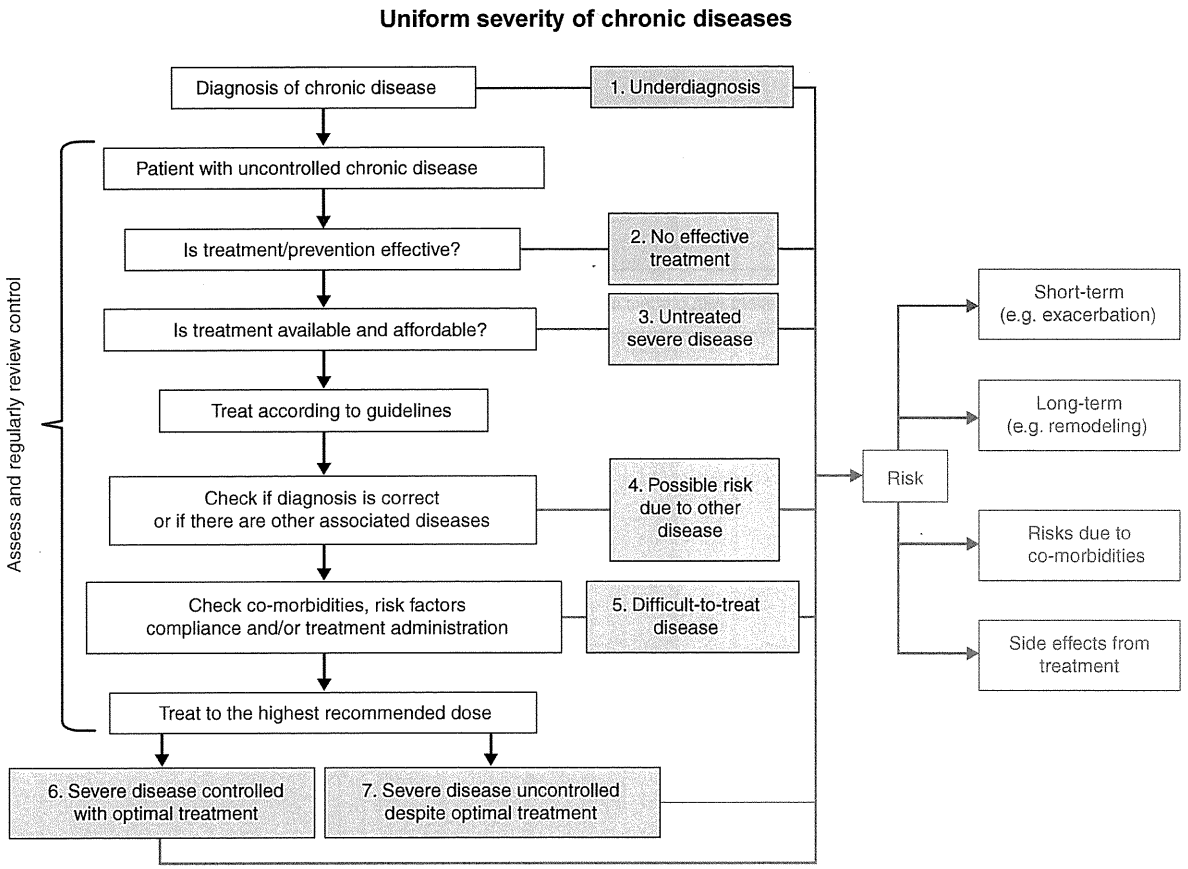


Figure 3. The concept of a uniform definition for NCD severity is based on control, responsiveness to treatment and risks (short, medium and long term). A single flow chart is proposed to define severity to improve phenotype characterization for all purposes (research, clinical and public health). It is based on diagnosis, therapeutic interventions and their availability/affordability, risk factors and co-morbidities.

multi-scale, multi-level systems approaches. To reduce their socio-economic and public health impacts, we propose that NCDs should be considered as the expression of a continuum or common group of diseases with intertwined gene-environment, socio-economic interactions and co-morbidities that lead to complex phenotypes specific for each individual. The 'systems medicine' concept, which takes a holistic view of health and disease, encapsulates this perspective. Systems medicine aims to tackle all components of the complexity of NCDs so as to understand these various phenotypes and hence enable prevention (Box 2), control through health promotion [50] and personalized medicine [51], and an efficient use of health service resources [52]. It does this through integrated care using multidisciplinary and teamwork approaches centered in primary and community care [53], including the essential ethical dimension.

Systems biology and medical informatics for P4 medicine of chronic NCDs

The main challenge regarding NCDs in the 21st century is to understand their complexity. Biology and medicine may be viewed as informational sciences requiring global systems methods using both hypothesis-driven and discovery-driven approaches. Systems medicine is the application of systems biology to medical research and practice [54,55]. Its objective is to integrate a variety of data at all relevant levels of cellular organization with clinical and patient-reported disease markers. It uses the power of computational and mathematical modeling to enable understanding of the mechanisms, prognosis, diagnosis and treatment of disease [56]. It involves a transition to predictive, preventive, personalized and participatory (P4) medicine, which is a shift from reactive to prospective medicine that extends far beyond what is usually covered by the term personalized medicine

Box 2: Glossary of terms

The classical definition of prevention [101] includes:

- **Primary prevention:** to avoid the development of disease.
- **Secondary prevention:** recognize a disease before it results in morbidity (or co-morbidity).
- **Tertiary prevention:** to reduce the negative impact of established disease by restoring function and reducing disease-related complications.

Expanding on the traditional model of prevention, Gordon [104] proposed a three-tiered preventative intervention classification system on the basis of the population for whom the measure is advisable based on a cost-benefit analysis:

- **Universal prevention** addresses the entire population (for example, national, local community, school, and district) and aims to prevent or delay risk factor exposure. All individuals, without screening, are provided with information and skills necessary to prevent the problem.
- **Selective prevention** focuses on groups whose risk of developing problems is above average. The subgroups may be distinguished by characteristics such as age, gender, family history, or economic status.
- **Indicated prevention** involves a screening process.

According to these definitions, **health promotion** [50] should be used for primary universal and selective prevention strategies, whereas **P4 medicine** (predictive, preventive, personalized and participatory) [51] should be used for primary, secondary and tertiary indicated prevention strategies.

Box 3: Key expected benefits of P4 medicine

To prevent the occurrence of NCDs by implementing effective action at societal and individual levels:

- To detect and diagnose disease at an early stage, when it can be controlled effectively.
- To stratify patients into groups, enabling the selection of optimal therapy.
- To reduce adverse drug reactions through the predictive or early assessment of individual drug responses and assessing genes leading to ineffective drug metabolism.
- To improve the selection of new biochemical targets for drug discovery.
- To reduce the time, cost, and failure rate of clinical trials for new therapies.
- To shift the emphasis in medicine from reaction to prevention and from disease to wellness.

[57,58]. It incorporates patient and population preferences for interventions and health states by implementing effective societal actions [57] with an important public health dimension [59]. It is likely to be the foundation of global health in the future (Box 3).

Thus, there is an urgent need for development of information management systems that can enable secure storage of heterogeneous data, including clinical data, and provide tools for the management, search and sharing of the data. Such information needs to accessible, shared between investigators, queried, and integrated in a controlled and secure manner with molecular profiles and images obtained from high-throughput facilities. For example, one prediction arising from considerations of the evolution of P4 medicine suggests that, in 10 years or so, each patient will be surrounded by a virtual cloud of billions of data points; we will need information technology to reduce this staggering data dimensionality to simple hypotheses about health and disease for each individual patient [57].

A systems biology approach that is unbiased by old classification systems can be used to find new biomarkers of co-morbidities, disease severity and progression. In this approach, phenotypes of NCDs are analyzed in an integrative manner using mathematical and statistical modeling, taking all diseases into account, and embedding co-morbidities, severity and follow-up of the patients through analyses in dynamic models (Figure 4). Unknown phenotypes are defined and further analyzed using iterative cycles of modeling and experimental testing. Novel biomarkers are identified combining datasets from genomics, epigenetics, proteomics, transcriptomics, metabolomics and metagenomics. These new complex biomarkers will need to be validated and replicated in independent controls or prospective patient cohorts [60]. Using methods used in non-medical complex model systems, it should be possible to monitor 'early warning signals', which predict the state of disease progression, and the occurrence of abrupt phase transitions (slowing down, increase in autocorrelation and variance) [61]. For example, in a mouse model of neurodegenerative disease, blood biomarkers have been shown to allow pre-symptomatic diagnosis and analysis of the stage of disease progression [62].

Modeling is a powerful tool for reducing the enormous complexity of comprehensive biological datasets to simple hypotheses. Modeling of the temporal behavior of disease read-outs at short [63] or long [64] intervals can identify sub-phenotypes of NCDs. Attempts to find novel biomarkers of disease development using a systems biology approach have been used to assess the mechanisms of severe asthma, allergy development [65] and cancer. One important role that biomarkers will have is to stratify a given disease into its different subtypes so that appropriate and distinct therapies can be selected for each subtype. Phenotypes can be modeled using statistical approaches, such as scale-free networks and Bayesian clustering models, that are based on the evaluation of NCDs as a whole, taking into account co-morbidities, severity and follow-up. This approach will

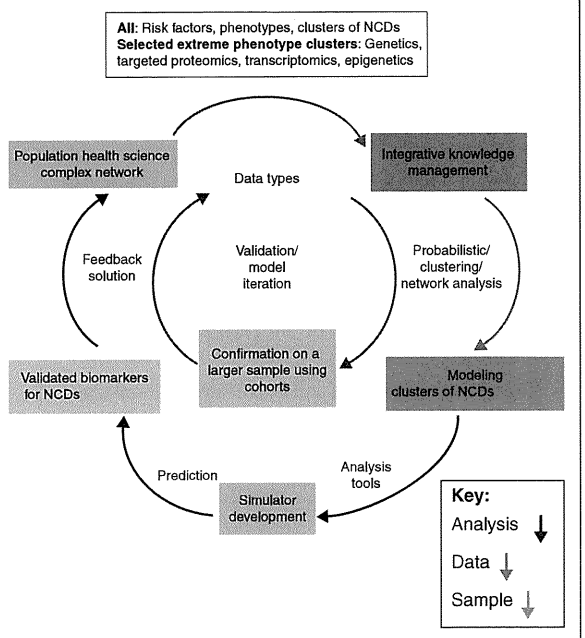


Figure 4. Iterative mathematical modeling to increase knowledge on NCDs. Various targeted or comprehensive data types are collected from samples of individuals with carefully defined phenotypes, processed using probabilistic and network analysis tools, and integrated into predictive models using a knowledge management and simulation platform, leading to the refinement of the classification of NCDs. Mechanistic hypotheses and complex biomarkers of NCDs generated through this process are then tested and validated iteratively using small then large cohorts of independent samples, providing potential diagnostic and therapeutic solutions for the general population.

make it possible to find intermediate phenotypes and patient-specific phenotypes. The challenge will be to develop efficient, automated and integrated workflows that predict the most suitable therapeutic strategy not only at the population level but, most importantly, at the individual patient level.

Bioinformatics, medical informatics and their interplay (sometimes termed biomedical informatics) will be key enablers in structuring, integrating and providing appropriate access to the enormous amount of relevant data and knowledge [66,67]. Medical informatics needs to provide ubiquitous and powerful electronic healthcare record technologies to securely aggregate and handle diverse, complex, and comprehensive data types [68]. Biomedical informatics must develop ways to use these content-rich electronic healthcare records to provide advanced decision support that considers all aspects of normal and disease biology, guided by clinically relevant insights and biomarker discovery research strategies [69,70]. Bioinformatics will need to constantly restructure and refine global data to distill the clinically useful elements and the derived models, so they can feed this information system in a real-time, automated fashion, constantly incorporating clinical expertise. P4 medicine is evolving so rapidly in its understanding of disease states that the individual patient's data must continually be re-examined so that new insights into the health and disease state of the individual can be gained. This general informatics framework, based on an advanced ICT infrastructure, will provide the basis for empowering P4 medicine.

Given the complexity of NCDs, bio-clinical scientific progress will depend critically on large-scale pooled analyses of high quality data from many biobanks [71] and bio-clinical studies (such as BioSHaRE-EU [72]). Biomedical informatics and knowledge management platforms have made significant advances towards enabling the development of technologies to organize molecular data at the level required for the complexity of NCD data [73,74]. Data analysis, integration and modeling require strict statistical procedures in order to avoid false discoveries [75]. They can be performed, for example, using the joint knowledge management platform of European Framework Program 7 (EU FP7) projects, including U-BIOPRED [76], MeDALL [65], AirPROM and SYNERGY-COPD, and using similar initiatives worldwide. Large-scale profiling to discover early markers of disease progression before the appearance of any symptoms has already been performed in a large prospective human cohort [77,78].

Complementary approaches using computational models that extend existing models derived from the Physiome project, including biomedical imaging, can be used together with statistical modeling of various types of clinical data to further define phenotypes and develop predictive models. These can be used within the framework of a fully integrated (preferably open source) knowledge management platform [79]. Such a platform for knowledge management, including annotation and ontologies, would then operate on top of the medical informatics infrastructure, setting the stage for a systems medicine approach to NCDs. In our collective experience these necessary aspects of medical informatics have a tendency to be overlooked in funding efforts targeting complex diseases.

Integrated care of chronic NCDs using P4 systems medicine

Integrated care, a core component of health and social care reforms, seeks to close the traditional gap between health and social care [80]. Population health sciences should integrate personalized medicine in public health interventions to prevent and manage NCDs in a cost-effective manner by involving all stakeholders, including patients [81]. The objectives of this proposed integration

are: (i) to investigate questions related to NCDs; (ii) to improve the quality of primary care; and (iii) to widely disseminate new information that will improve overall health at both a local and national level [82]. Chronic diseases can disconnect individuals from their usual milieu, with negative implications for physical, social and mental well-being. Moving beyond the disease-by-disease approach to tackle NCDs demands an improved understanding of NCD by patients, and a better understanding of their common causes. At the local level, strategies such as community oriented primary care can link and reinforce personal and public health efforts [83].

To understand, preserve and improve the health of human populations and individuals, an integrated research strategy should include all components of research on NCDs and be integrated for optimal patient management [84,85]. Careful evaluation is needed of: (i) the acceptance of multi-morbidity of NCDs by the patient, with particular attention to cultural and social barriers, gender and age; (ii) the engagement of patients in decisions regarding management [86], research and clinical trials [55,57]; and (iii) the improvement of quality of life that would result from the proposed management. Targeting NCDs and their comorbidities will directly affect healthy aging, which has been described as a 'keystone for a sustainable Europe' [87]. Screening, early diagnosis, prevention and treatment of hidden comorbidities in patients with diagnosed NCDs will reduce their morbidity and increase healthy life years.

The direct and indirect costs of uncontrolled NCDs are substantial for the patient, the family and society, especially in underserved populations [9]. P4 medicine should be put into the context of health economics to show that expensive strategies are cost-effective [55,57]. Chronic diseases place a considerable economic burden on the society and increase inequities. The social dimension of NCDs needs to be pursued in the economic and employment fields. The net social benefit of improving medical and social care related to NCDs should take co-benefits into account. Health costs for NCDs should be balanced with health benefits, wealth creation and economic development. The management of NCDs requires the coordination of stakeholders in the public and private sectors within a governance framework that includes networks of care. Therefore, research should be done to identify social determinants and to create public health systems that translate efficacy into effectiveness in the community [88]. Moreover, strengthening health equity across nations and socioeconomic groups is needed to meet the ambitions of the Commission on Social Determinants of Health, who have proposed closing the health gap between nations and groups in a generation [89].

Values are the basis of most actions in health and the economy, and these values are often not made explicit. Changing paradigms and approaches to NCDs may challenge fundamental societal values and professional habits [59,90]. The apparent contradiction between the development of a more tailored medical approach to NCDs and the public health dimensions of their prevention and care needs to be addressed using a value-based analysis. Thus, a thorough analysis of values underlying P4 medicine should be conducted in diverse contexts and should become part of the basis of decision-making. The respective weight of the multiple stakeholders involved in the priority setting must be made clear, with transparency and proportionality as key features. P4 medicine development should be a global aim and not a privilege of 'rich' countries. Using data obtained from all components of research, guidelines on NCDs applicable to primary care could be developed using up-to-date methodology [91,92]. Policies for implementation could then be proposed, to translate the concept of NCD into practice. They should distribute the burdens equitably, also considering gender and age.

Multidisciplinary training of all stakeholders, with particular emphasis on the participation of patient associations, is a further essential component. Many health and non-health professionals need to be educated in the general approach to the research and management of patients with NCDs. Innovative training programs using ICT will be essential in this implementation. Such education will also need to address questions of how to teach the subject and how people learn it, rather than merely regarding education as a process of transmission and transaction for everyone involved. This includes taking into account points of view, habits of mind, and all the information requested for the needs of the strategy. The educational program needs to forge educational systems to help participants think in a coherent way about NCDs. A module of the program should be developed around patient feedback to help them be engaged in all aspects of NCDs, including research.

Many patients with NCDs live in developing countries where medications and services are often unavailable or inaccessible. Effective medications, such as inhaled corticosteroids for asthma [93] or insulin for diabetes, should be made available for all patients [94]. In addition, there should be a global cost-effective application of P4 medicine across the world [95]. It is likely that genomic applications and ICT will become available to many developing countries at a relatively low cost in the next few years. In addition, new private-public strategic partnerships, such as the pre-competitive Innovative Medicines Initiative, a joint undertaking of the European Union and the European Federation of Pharmaceutical Industry Associations [96], and the Program on

Public-Private Partnerships of the United States National Institutes of Health Roadmap [97], are required to overcome the bottlenecks in the development of new treatment strategies [98]. WHO actively supports capacity building, especially in developing countries, fosters partnerships around the world, and works to narrow the gap in healthcare inequities through access to innovative approaches that take into account different health systems, economic and cultural factors. Despite the growing consensus for the need for health system strengthening, there is little agreement on strategies for its implementation [99]. Widely accepted guiding principles should be developed with a common language for strategy development and communication for the global community in general [100] and for NCDs in particular.

Conclusions

NCD management needs to move towards integrated care, global strategies and multi-modal systems approaches, which will reduce the burden and societal impact of NCDs. To this end, we propose that NCDs must be considered as the expression of a common group of diseases with different risk factors, socio-economic determinants and co-morbidities. This will enable the application of P4 medicine principles to NCDs, exploiting their commonalities, bringing improved global healthcare and the reduction of inequities around the world. The expected results targeted to better support for patients include: (i) better structuring of translational research and development for NCDs; (ii) greatly enhanced prevention and treatment capabilities; (iii) innovative healthcare systems with implementation of follow-up procedures directly in the homes of patients; (iv) slowing down of health expenditure increase; and (v) new interdisciplinary training curricula.

Abbreviations

AIRPROM, AIRway disease, PRedicting Outcomes through patient specific computational Modeling (FP7); BioSHare-EU, Biobank Standardization and Harmonization for Research Excellence in the European Union (FP7); ICT, information communication technology; MeDALL, Mechanisms of the Development of ALLergy (FP7); NAEPP-EPR3, National Asthma Education and Prevention Program, Expert Report 3; NCD, non-communicable disease; P4, predictive, preventive, personalized and participatory; U-BIOPRED, Unbiased BIOmarkers in PREDiction of respiratory disease outcomes (FP7); UN, United Nations; WHO, World Health Organization.

Competing interests

The authors declare that they have no competing interests in relation to the content of this article.

Acknowledgements

Part of the conceptual work presented has received support from the European commission FP7 projects AIRProm (Grant Agreement FP7 270194), BioSHaRE-EU (Grant Agreement FP7 261433), MeDALL (Grant Agreement FP7 264357), SYNERGY-COPD (Grant Agreement) and U-BIOPRED (Grant Agreement IMI 115010). JB, JMA, AC-T, FK, MLK, SP, CP and CA were supported by MeDALL; PJS, IMA and KFC were supported by U-BIOPRED; JR and AA were supported by SYNERGY-COPD; CB was supported by AIRProm; AC-T was supported by BioSHaRE-EU.

The positions, proposals and ideas expressed in this paper have been discussed by several authors (CA, ZC, LH, AB, JB, AC, SA, DC, DN) during the inaugural event of the European Institute for Systems Biology and Medicine of the Systemoscope International Consortium at the Biovision World Life Sciences Forum in Lyon on 28 March 2011.

Author details

[1]Department of Respiratory Diseases, Arnaud de Villeneuve Hospital, CHU Montpellier, INSERM CESP U1018, Villejuif, France. [2]Centre for Research in Environmental Epidemiology, Municipal Institute of Medical Research, Epidemiologia y Salud Publica, Universitat Pompeu Fabra, Doctor Aiguader, 88, E-08003 Barcelona, Spain. [3]Academic Medical Centre, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands. [4]Cellular and Molecular Biology, Imperial College, South Kensington Campus, London SW7 2AZ, UK. [5]National Heart and Lung Institute, Imperial College, South Kensington Campus, London SW7 2AZ, UK. [6]Institut Clínic del Tòrax, Hospital Clínic, IDIBAPS, CIBERES, Universitat de Barcelona, Spain. [7]Department of Infection, Immunity and Inflammation, University of Leicester, Sciences Building, University Road, Leicester, LE1 9HN, UK. [8]Epidemiology, Public Health, Risks, Chronic Diseases and Handicap, INSERM U558, Toulouse, France. [9]IRCCS San Raffaele, Via della Pisana, 235, Rome, Italy. [10]Institut Pasteur, Bab Bhar, Avenue Jugurtha, Tunis, 71 843 755, Tunisia. [11]Division of Medical Genetics, University of Geneva Medical School, 1 rue Michel-Servet, 1211 Geneva 4, Switzerland. [12]Department of Diabetology, Montpellier, France. [13]Telethon Institute of Genomics and Medicine, Via Pietro Castellino, 111 80131 – Napoli, Italy. [14]Department of Pediatrics, University of Padova, Padova, Giustiniani, 3 – 35128, Italy. [15]Scientific Centre of Children's Health, Russian Academy of Medical Sciences, Lomonosovskiy prospect, 2/62, 117963,Moscow, Russia. [16]Department of Dermatology and Allergy, University of Bonn, Sigmund-Freud-Str. 25, 53105 Bonn, Germany. [17]Institut de Génomique Fonctionnelle, CNRS, UMR 5203, INSERM, U661, Université Montpellier 1 and 2, Montpellier, France. [18]Institute of Genomics and Integrative Biology, Near Jubilee Hall, Mall Road, Delhi-110 007, New Delhi, India. [19]Pulmonary Division, Albert Michallon University Hospital, Albert Bonniot Cancer Research Institute, La Tronche, Grenoble, France. [20]Endocrine Diseases, Lapeyronie Hospital, Montpellier, France. [21] Department of Physiology, Nîmes University Hospital, Place du Professeur Robert Debré. 30029 Nîmes Cedex 9, France [22]Department of Microbiology, Tumour and Cell Biology, Karolinska Institute, Nobels väg 16, KI Solna Campus, Box 280, SE-171 77 Stockholm, Sweden. [23]Department of Medical and Surgical Specialties, University of Modena and Regio Emilia, Modena, Italy. [24]Imperial College London, London, UK. [25]Institute for Systems Biology, Seattle, 401 Terry Avenue, North Seattle, WA 98109-5234, USA. [26]National Institute of Genetics, Mishima, Japan. [27]Auckland Bioengineering Institute, University of Auckland, Level 6, 70 Symonds Street Auckland, 1010. New Zealand. [28]Clinical Unit for Osteoarticular Diseases, and INSERM U844, Montpellier, France. [29]Centre for Research in Epidemiology and Population Health, INSERM U1018, Villejuif, France. [30]Singapore Immunology Network, 8A Biomedical Grove, Level 4 Immunos Building, 138648 Singapore . [31]Medical University of Lodz, Poland. [32]Department of Molecular Genetics, Weizmann Institute of Science, P.O. Box 26 Rehovot 76100, Israel. [33]Health Economy and Management, Paris-Dauphine University, Paris, France. [34]Biotechnology and Biotherapy, IRCM, Paris, France. [35]Department of Medicine, Groote Schuur Hospital and University of Cape Town, South Africa. [36]Department of Physiology, Montpellier University, and INSERM U1046, France. [37]The Estonian Genome Center of University of Tartu, Tartu, Estonia. [38]Epsylon, Montpellier, France. [39]Department of Physiology, University of Oxford, Le Gros Clark Building, South Parks Road, Oxford OX1 3QX, UK. [40]Department of Molecular Biology and Genetics, Bilkent University, Faculty of Science, B Building, 06800 Ankara, Turkey. [41]European Patient's Forum (EPF) and European Federation of Allergy and Airways Diseases Patients Associations (EFA), Brussels, Belgium. [42]Department of Medicine, University of Kiel, Germany. [43]Department of Family Medicine, University of Wisconsin, 1100 Delaplaine Ct. Madison, WI 53715-1896, USA. [44]Department of Public Health, AL. JEROZOLIMSKIE 87, 02-001 Warsaw, Poland. [45]Departments of Clinical Epidemiology and Biostatistics and of Medicine, McMaster University, 1280 Main Street West, Rm. 2C12, L8S 4K1 Hamilton, ON, Canada. [46]Institute of Pharmacology, University of Bern, Friedbühlstrasse 49, CH-3010 Bern, Switzerland. [47]Cancer Biology and Epigenomics Program, Children's Memorial Research Center and Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, USA. [48]Research Centre for Molecular Medicine, Lazarettgasse 14, AKH BT 25.3, A-1090, Vienna, Austria. [49]Department of Medicine, Karolinska Institute, Solna,