Full Genotyping of a Highly Polymorphic Human Gene Trait by Time-Resolved Fluorescence Resonance Energy Transfer

Edoardo Totè¹, Marco Lamperti¹, Maria Bondani², Domenico Salerno³, Valeria Cassina³, Luca Nardo³*

¹Department of Science and High Technology, University of Insubria, Como, Italy, ²C. N. R. (Consiglio Nazionale delle Ricerche), Institute for Photonics and Nanotechnology, Como, Italy, ³Department of Health Sciences, University of Milan Bicocca, Monza, Italy

Abstract
The ability of detecting the subtle variations occurring, among different individuals, within specific DNA sequences encompassed in highly polymorphic genes discloses new applications in genomics and diagnostics. DQB1 is a gene of the HLA-II DQ locus of the Human Leukocyte Antigens (HLA) system. The polymorphisms of the trait of the DQB1 gene including codons 52–57 modulate the susceptibility to a number of severe pathologies. Moreover, the donor-receiver tissue compatibility in bone marrow transplantations is routinely assessed through crossed genotyping of DQB and DQA. For the above reasons, the development of rapid, reliable and cost-effective typing technologies of DQB1 in general, and more specifically of the codons 52–57, is a relevant although challenging task. Quantitative assessment of the fluorescence resonance energy transfer (FRET) efficiency between chromophores labelling the opposite ends of gene-specific oligonucleotide probes has proven to be a powerful tool to type DNA polymorphisms with single-nucleotide resolution. The FRET efficiency can be most conveniently quantified by applying a time-resolved fluorescence analysis methodology, i.e. time-correlated single-photon counting, which allows working on very diluted template specimens and in the presence of fluorescent contaminants. Here we present a full in-vitro characterization of the fluorescence responses of two probes when hybridized to oligonucleotide mixtures mimicking all the possible genotypes of the codons 52–57 trait of DQB1 (8 homozygous and 28 heterozygous). We show that each genotype can be effectively tagged by the combination of the fluorescence decay constants extrapolated from the data obtained with such probes.

Introduction
The completion of the Human Genome Project has evidenced a large number of subtle variations (polymorphisms) among the genomes of different individuals of the same population [1]. Gene polymorphism is responsible for the expression of many variants of a protein among different subjects of the same species. To be classified into polymorphism, variations of gene sequences must appear in at least 1% of the population. In the latter case, the genetic variants are called alleles. The great majority of DNA polymorphisms consist in single-base variations, i.e. single-nucleotide polymorphisms (SNPs) [1]. Most SNPs occur in introns, thus they are silent, i.e. they have no effect on the individual phenotypes. However, SNPs, though rarely being directly pathogenic, are frequently correlated to the occurrence of multigenic diseases [2]. Moreover, certain SNPs result in mutations of genes that are pivotal in tumorigenesis. An example is the tumor suppressor gene p53 [3]. More recently, SNPs have been also exploited as markers in pharmacogenomics [4]. For all the above reasons, the possibility to perform wide range population screening of the relevant polymorphic genes would provide a powerful tool for disease prediction.

A particularly interesting highly polymorphic locus within the human genome is the major histocompatibility complex, located in the short arm of chromosome 6. It comprises many genes coding for the human leukocyte antigen (HLA) proteins that, originally discovered for their role in the rejection of transplanted organs, play a crucial role in the triggering of the immune response against pathogens. Indeed, they guide the development and the activation of T lymphocytes. There are two types of HLA molecules, namely class I and class II, that activate functionally distinct T cell populations, CD8+ T cells and CD4+ T cells, respectively [5]. The expression of certain HLA molecules has been associated with the susceptibility to several autoimmune diseases, including celiac disease [6–9], multiple sclerosis [10,11], Hashimoto thyroiditis [12], Graves disease [12], narcolepsy [13], selective immunoglobulin A (IgA) as well as common variable immunodeficiencies (CVI) [14], sarcoidosis [15], primary biliary cirrhosis [16], autoimmune [17] and viral [18] hepatitis, insulin-dependent diabetes mellitus (IDDM) [12,19,20]. For all the above pathologies, genetic markers are found in the DQB1 gene, namely in the short though highly polymorphic trait encompassing codons 52–57. The latter trait is manifested in eight allelic variants: DQB1-

Citation: Totè E, Lamperti M, Bondani M, Salerno D, Cassina V, et al. (2014) Full Genotyping of a Highly Polymorphic Human Gene Trait by Time-Resolved Fluorescence Resonance Energy Transfer. PLoS ONE 9(9): e107310. doi:10.1371/journal.pone.0107310

Editor: Sabato D’Auria, CNR, Italy
Received July 10, 2014; Accepted August 10, 2014; Published September 12, 2014

Competing Interests: The authors have declared that no competing interests exist.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: The authors have no funding or support to report.
Table 1. Compendium of disease susceptibility/protection patterns correlated with DQB1 genotype.

| Disease                  | Predisposing genotypes                      | Protective genotypes                  | Notes                                      |
|--------------------------|---------------------------------------------|---------------------------------------|--------------------------------------------|
| Celiac disease           | 0201/0XXX; 0302/0XXX; 0302/0301; 0302/0301 | –                                     | 90% patients bear the 0201/0201 genotype  |
| IDDM                     | 0201/0XXX; 0302/0XXX                         | 0602/0XXX; 0301/0XXX                  | –                                          |
| IGA-CVID                 | Non-Asp 57 in both alleles                   | Asp 57 in both alleles                | 0201/0201 genotype is maximally predisposing|
| Graves disease           | 0201/0XXX                                   | –                                     | 0201/0302 genotype is maximally predisposing|
| Hashimoto thyroiditis    | 0201/0XXX                                   | 0302/0XXX                             | –                                          |
| Primary biliary cirrhosis| 0402/0402                                   | –                                     | –                                          |
| Sarcoidosis              | 0602/0XXX                                   | 0201/0XXX                             | –                                          |
| Narcolepsy               | 0602/0XXX                                   | –                                     | 0602/0602 genotype is maximally predisposing|
| Type 1 autoimmune hepatitis | –                                      | 0301/0XXX                             | –                                          |
| Viral hepatitis C         | –                                          | 0502/0XXX                             | –                                          |
to the relevant fluorescence signal emitted by the probe of an unpredictable amount of fluorescence produced by endogenous fluorophores. Finally, detectors endowed with light-intensity-proportional responses must be used, whose low quantum efficiency makes PCR amplification hardly avoidable. The above-mentioned drawbacks can be overcome by quantitatively analyzing the variations in FRET efficiency through measurements of the fluorescence decay time $\tau_D$ of $D$. Indeed, $\tau_D$ can be determined by means of time-correlated single-photon counting (TCSPC). TCSPC works by detecting and timing one photon per excitation pulse at most. The subsequent reconstruction of the statistical distribution of the detection times relative to the excitation pulses yields the fluorescence decay pattern, whose fitting provides the value of $\tau_D$. For this reason, TCSPC is intrinsically independent from the D concentration and takes advantage of the usage of Geiger-mode light detectors, such as single-photon avalanche diodes, which are sensitive to even extremely dim fluorescence pulses, virtually reduced to few/single photons.

We recently showed that by this method it is possible to pursue the recognition of the DQB1-0201 genotype, conferring susceptibility to, e.g., celiac disease and IDDM, by working on template DNA contained in untreated (i.e. unamplified and non-purified) cell extracts [27]. To this aim we designed an oligonucleotide probe dual labelled at its 5’ and 3’ ends, respectively, with a fluorescent dye and with a non-emitting FRET acceptor, and with sequence complementary to the region of DQB1-0201 allele encompassing codons 51–57 and the first nucleotide of codon 58. Namely, the fluorophore was tetramethyl-rhodamine (TAMRA) while the acceptor was black-hole quencher 2 (BHQ2). The latter probe, hereafter Probe1, was made to hybridized to oligonucleotides mimicking any of the eight possible allelic variants of that region of DQB1. The $\tau_D$ of TAMRA attached to Probe1 annealed to each of the target oligonucleotides was measured with 30 ps resolution. Because each target oligonucleotide was characterized by distinctive mismatches with respect to Probe1, we expected the resulting probe/target duplexes to have slightly different conformations, hopefully reflected in significantly different $\tau_D$ values. Indeed, the allele of interest, DQB1-0201, yielded a $\tau_D$ value distinct from the most similar of the others for as much as 42 standard deviations. Moreover, we managed to reveal the presence of the DQB1-0201 allelic variant also in heterozygous conditions, and to discriminate the latter from the homozygous case. Furthermore, Probe1 demonstrated to be sufficiently selective to specifically recognize the DQB1 trait of interest within the whole genome. Finally, we were able to replicate the $\tau_D$ results by using unpurified template DNA contained in bare cell lysates. However, we failed to obtain a full discrimination of alleles. Namely, even in the preliminary calibration measurements undergone on the oligonucleotide targets, Probe1 yielded hardly distinguishable $\tau_D$ values when annealed to DQB1-0503, DQB1-0602, and DQB1-0402-like sequences. With the goal of breaking such a degeneracy on $\tau_D$ values and pursuing a complete genotyping of the DQB1 trait of interest, more recently we performed additional TCSPC experiments aimed at characterizing the fluorescence decay of TAMRA attached to DNA duplexes obtained by annealing the same target oligonucleotides of above with a second oligonucleotide probe (Probe2), complementary to DQB1-0503 [28]. We obtained notably different $\tau_D$ values for Probe2 hybrids with DQB1-05031, DQB1-0602, and DQB1-0402-like targets.

In this work we present a systematic study of the fluorescence decay patterns produced by Probe1 and Probe2 upon reaction with oligonucleotide samples mimicking all the possible genotypes of the DQB1 trait encompassing codons 52–57 (i.e. 8 samples containing a single target oligonucleotide variant and the 28 samples resulting by mixing 1:1 concentrations of any possible pair of oligonucleotides). We demonstrate that, while the “homozygous” samples can be discriminated by cross-combining the $\tau_D$ value obtained with Probe1 with that obtained with Probe2, for the “heterozygous” ones it is still possible to extract an average decay time value from the decay data obtained with both probes, $<\tau_D>$. The $(<\tau_D_{\text{Probe1}}>; <\tau_D_{\text{Probe2}}>)$ pair are an exhaustive tag-parameter allowing to discriminate among the different heterozygous genotypes.

Materials and Methods

Oligonucleotides and Sample Preparation

In Figure 1 the eight target oligonucleotides reproducing the possible sequences of the 22-bases polymorphic trait encompassing codons 51 to 57 and the first base of codon 58 are depicted. In the same figure, the TAMRA-BHQ2 dual-labelled probes Probe1 and Probe2, with sequences complementary to the DQB1-0201 and DQB1-0503 allele-mimicking oligos, respectively, are also shown. Mismatch-sites of the target oligonucleotides with respect to Probe1 are black-shaded, while those with respect to Probe2 are light-gray shaded. Mismatching sites common to both probes are indicated by dark-grey shading. The target oligonucleotides were purchased by TIB Molbiol (Genova, Italy) and were desalted and further purified by double reverse-phase procedure. The dual-labelled probes were purchased by Eurofins MWG Operon (Ebersberg, Germany) and purified by high-performance liquid chromatography. A nominal amount of 20 nanomoles of the lyophilized oligonucleotides was resuspended in appropriate volumes (indicated by the suppliers) of Tris-HCl EDTA buffer at pH 7.6 and 10 mM ionic strength, to obtain 50 µM concentrated solutions. Equal amounts (10 µL) of probe and target stock solutions were mixed to obtain the “homozygous” solutions. To attain the solutions mimicking the heterozygous genotypes, e.g. DQB1-0XXX/DQB1-0YYY, 5 µL of DQB1-0XXX target oligo and 5 µL of DQB1-0YYY were mixed to 10 µL of probe. The annealing protocol, which is fully described elsewhere [27], consisted in heating the samples to 98°C in order to denature any residual secondary structures of the single-stranded oligonucleotides, cooling at the rate of 1°C/min to the annealing temperature and allowing the probe to react with the complementary oligonucleotide for 10 minutes. For “heterozygous” samples, the cooling was stopped for 10 minutes at the annealing temperatures of both targets. Finally, the solutions were cooled to 25°C at 3°C/min rate. The thermostat of a Thermofast Finnigan gas chromatograph (San Jose, CA, USA) was used as the temperature controller. The obtained solutions were diluted in phosphate buffer at pH 7.6 and 150 mM ionic strength to a final probe concentration of 250 nM, to produce the sample for TCSPC analysis.

Time-Correlated Single-Photon Counting Apparatus

The fluorescence excitation/detection apparatus is extensively described elsewhere [26]. Briefly, a 113 MHz repetition rate and a train of excitation pulses (pulse duration 6.4 ps) at 532 nm is provided by a Nd:VAN cw-mode-locked laser (GE-100, Time Bandwidth Products, Zurich, CH) and attenuated by a factor of 10^5 by means of neutral density filters, in order to achieve single-photon regime (photon counts rate <100 kHz for all samples). Fluorescence is collected at 90° to the excitation beam through a long-wavelength-pass filter with cut at 550 nm, and focused onto the sensitive area of a single-photon avalanche diode (PD5M5, Micro-photon devices, Bolzano, IT) by a 20X microscope objective. The avalanche pulses give the START signal to an integrated TCSPC.
be able to bypass deformations induced in the decay histograms by assuring to start the fitting far enough from excitation in order to channels (i.e., 488 ps, thus 16 IRF full-width at half maximum) IRF deconvolution algorithm was applied. However, the fitting (IRF, decay constants, were used to tag each genotype. In both cases, decays. The extracted heterozygous decay distributions were fitted to single exponential the results of the above fits was attempted. Secondly, the "clean" heterozygous data ideally represent the decay distributions built up uniquely by the fluorescence photons emitted by TAMRA molecules labeling intact (i.e. BHQ2-functionalized) probes correctly hybridized to a target oligonucleotide. Two different approaches were applied to analyze such decays. Firstly, two-exponential fits were performed, with time constants fixed at any possible pair of \( t_D \) values measured for homozygous alleles.

Data Analysis

The decay patterns of all the probe/target duplex samples and of the single-stranded probe were measured for 5 minutes. The data processing protocol devised in order to get read of the contributions to the fluorescence decays of unquenched TAMRA molecules and non-hybridized (i.e. single-stranded) probe molecules, which is detailed in previous works [25–27] and briefly summarized in Fig. 2, was standardized and automated by compiling a Matlab routine. We further subtracted a constant background value accounting for non-time-correlated environmental light, which was evaluated by acquiring a photon lag-times pattern with the sample substituted by a water-filled cell for the same measuring time, and averaging the counts values over 1000 channels far from the pulse response peak. Preliminary to the new experiments on heterozygous genotypes mimicking samples, we repeated (in one parallel only) measurements on homozygous genotypes mimicking samples. The so obtained data, after being "cleaned" by using the Matlab routine, were fitted to single exponential decays such as in previous works [27,28].

The "clean" heterozygous data ideally represent the decay distributions built up uniquely by the fluorescence photons emitted by TAMRA molecules labelling intact (i.e. BHQ2-functionalized) probes correctly hybridized to a target oligonucleotide. Two different approaches were applied to analyze such decays. Firstly, two-exponential fits were performed, with time constants fixed at any possible pair of \( t_D \) values measured for homozygous alleles. Ex-post discrimination of the correct genotype by comparison of the results of the above fits was attempted. Secondly, the "clean" heterozygous decay distributions were fitted to single exponential decays. The extracted \( t_D \) values, representing a sort of average decay constants, were used to tag each genotype. In both cases, due to the negligible spread of the instrumental response function (IRF, \(<30\) ps) with respect to the measured \( t_D \) values \((>2\) ns), no IRF deconvolution algorithm was applied. However, the fitting interval was chosen to be of 2800 channels, starting from 200 channels (i.e., 488 ps, thus 16 IRF full-width at half maximum) after the detected excitation pulse peak. This choice, besides assuring to start the fitting far enough from excitation in order to be able to bypass deformations induced in the decay histograms by convolution with the IRF, allowed to exclude the board non-linearity region (i.e., the transient in which the current flowing through the board time-to-amplitude converter switches on and off).

It should be mentioned that putative systematic errors might arise from an incorrect choice of the fitting interval. However, fitting parameters values stable within few percent were obtained for exemplary datasets upon variation of the initial channel of the fitting interval in the range from 170 to 600 channels after the detected excitation pulse peak, with the last channel fixed at 3000 channels after excitation. Moreover, because our measurements are differential in nature and the \( t_D \) values are simply used as tags of the corresponding genotypes throughout the manuscript, without attempting any absolute evaluation of the donor-to-acceptor distance, systematic errors of any source would not invalidate our conclusions.

All the experiments on heterozygous samples were performed in triplicate, and the values reported below are the average on the three parallel samples. The pertaining standard deviations were used as experimental errors. The built-in Matlab fitting tool which we exploited within our routine automatically yields confidence intervals for the fitting parameters values, which were in all cases smaller than such standard deviations.

Results and Discussion

Firstly, we shortly report on the results yielded by analysis of the repeated experiments on homozygous samples. The \( t_D \) values obtained from the new datasets were equal, within the experimental errors, to those published in the past. This stems in support of both the repeatability and reliability of the typing protocol and the correctness of the new automatized data processing routine. The \( t_D \) values resulting by averaging the old with the new fitting parameters, with errors given by standard deviations, are summarized in Table 2. The new data are reported in parenthesis, while the old ones are published in [28].

We now come to describe in details the results obtained on the heterozygous genotype mimicking samples. Theoretically, for both probes and for any given "heterozygous" sample DQB1-0XXX/DQB1-0YYY, after subtraction of the unquenched TAMRA and non-hybridized probe transients and of the constant background, there remain only two exponential components building up the "cleaned" fluorescence decay. The first component is produced by the part of the probe annealed to the target oligonucleotide bearing the sequence of the DQB1-0XXX allele. This should decay with the time constant \( t_D_{XXX} \), tagging the homozygous DQB1-0XXX genotype, reported in Table 2. The other compo-
Figure 2. Main steps of the data cleaning procedure shown for an exemplary dataset, namely one of those obtained with the DQB1-0201/0301 genotype mimicking oligonucleotide mixture annealed to Probe2. Upper panel: the decay patterns of the single-stranded probe and of the duplex are synchronized and the probe decay is fitted to a three-exponential-components model. Central panel: the longest lived component of the fit, which is attributed to non-quenched TAMRA emission, is subtracted to both the probe and the duplex distributions, which are then normalized to their integrals and fitted to a 2-exponential and a 3-exponential model, respectively. Lower panel: the weighted subtraction of the non-hybridized probe fraction to the duplex decay yields the net decay of TAMRA molecules attached to intact (i.e. bearing also BHQ2) and correctly hybridized probes. The latter can either be interpolated with a single-exponential or fitted to a constrained 2-exponential model in which the decay time constants values are fixed to those measured for any pair of alleles, the relative amplitudes are allowed to vary and their values analyzed in order to derive the correct genotype (see text for details). In the figure the 2-exponential fitting curve obtained by using the decay time constants pertaining to the correct pair of alleles is reported. doi:10.1371/journal.pone.0107310.g002
values obviously do not allow unambiguous attribution of genotype is not feasible by means of this analysis, which was thus abandoned.

Interestingly, the average decay time values \( <\tau_D> \) (see fifth column in Tables 3 and 4) calculated from any of the surviving set of fitting parameters as \( <\tau_D> = (\tau_{D,XXX} + \tau_{D,YYY})/2 \) are very similar, and equal within the experimental errors to the \( <\tau_D> \) values yielded by both single exponential fit \( 2566\pm9 \) ps and \( 2279\pm11 \) ps for the DQB1-0201/DQB1-0502 sample and Probe1 and Probe2, respectively) and unconstrained double exponential fits. This fact suggests that fitting convergence is mainly dictated by retrieval of the correct average decay time value. Hence, we decided to use the \( <\tau_D> \) value restituted by single exponential fit as the representative tag-parameter also for the heterozygous sample, in spite of the somewhat lower quality of fits, which in turn leads to higher uncertainties on the average \( <\tau_D> \) values calculated over the three parallel samples. The results of such single-exponential fits are reported in Table 5.

We first comment on the results obtained with Probe1. Although the \( <\tau_D> \) values obviously do not allow unambiguous discrimination of the various heterozygous genotypes (the latter was not possible with this probe even limiting the analysis to the homozygous genotypes), relevant information concerning the Asp/non-Asp nature of residue 57 can be extracted from these data. In Fig. 3 we show the dendrogram resulting from cluster analysis of the Probe1 data obtained assuming Euclidean metric.

\[ \Delta \tau_j \equiv \langle \tau_D \rangle - <\tau_D> \] as the distance and applying the complete-linkage cluster joining criterion. In the latter, the distance between two clusters is defined as the maximum distance between one element belonging to the first and one element belonging to the second cluster. The heterozygous genotypes bearing both alleles coding for Asp residues cluster up in a common family (red branches of the dendrogram). Even more relevantly, the genotype associated to maximum proneness to Graves disease and to incremented proneness to celiac disease, i.e. DQB1-0201/DQB1-0302, clusters on its own (black branch). In addition, all and only the heterozygous genotypes conveying \( per-se \) susceptibility to celiac disease (i.e. DQB1-0201/DQB1-XXX and DQB1-0501/DQB1-0302) make up the second (green) family of the dendrogram. It is also worth noting that the genotypes containing the maximally disease-correlated allele DQB1-0201 (conferring susceptibility to celiac and Graves diseases, IDDM, Hashimoto thyroiditis, IGA, and CVID, and protection with respect to sarcoidosis) all cluster in the above two families. Finally, the blue cluster contains all the DQB1-0302/DQB1-0XXX genotypes different from DQB1-0201/DQB1-0302, and only one additional “extraneous” genotype, namely DQB1-0501/ DQB1-0301. We recall that the presence of the DQB1-0302 allele is correlated to increased risk factor for IDDM, IGA and CVID and to protection with respect to Hashimoto thyroiditis. We also attempted cluster analysis of the Probe2 \( <\tau_D> \) values, but the results (reported in the dendrogram of Fig. 4) are less suggestive from a diagnostic standpoint, as the segregated families do not contain genotypes endowing with susceptibility to a common disease. Moreover, also with this probe unambiguous recognition of any possible genotype is not feasible (e.g. the DQB1-0501/ DQB1-0302 genotype is not distinguishable from DQB1-0602/ DQB1-0302, DQB1-0201/DQB1-0501 from DQB1-0201/ DQB1-0602, DQB1-0201/DQB1-0402 from DQB1-0301/ DQB1-0503, and DQB1-0501/DQB1-0402 from DQB1-0301/ DQB1-0502, as evidenced by the negligible distance values visualized in the dendrogram).

### Table 2. Fluorescence decay time, \( \tau_D \), of the TAMRA fluorophore linked to Probe1 and Probe2 hybridized to oligonucleotides mimicking all the allelic variants of the DQ\( B1 \) codon 51–57 trait (cumulative average of previously published data [28] with those derived in the check experiments presented here, which are reported in parenthesis.

| Genotype       | \( f_{xx} \times 10^4 \) | \( f_{yy} \times 10^4 \) | \( \chi^2 \) | \( <\tau_D> \) | \( F_{xxx} \) |
|----------------|--------------------------|--------------------------|-------------|--------------|------------|
| 0201/0502      | 3.1425                   | 2.8774                   | 0.9791      | 2571.57      | 0.52       |
| 0302/0602      | 5.0585                   | 0.9528                   | 0.9806      | 2572.21      | 0.84       |
| 0302/0301      | 4.6532                   | 1.3574                   | 0.9807      | 2572.35      | 0.77       |
| 0302/0501      | 4.1317                   | 1.8785                   | 0.9808      | 2572.43      | 0.69       |

In the last two columns the calculated values of the average lifetime and fraction of probe hybridized to the DQB1-0XXX allele are reported.

### Table 3. Initial amplitudes and \( \chi^2 \) values yielded by the fitting of the Probe1 data for the genotypes surviving the constrained fit analysis with both probes.

| Genotype       | \( f_{xx} \times 10^4 \) | \( f_{yy} \times 10^4 \) | \( \chi^2 \) | \( <\tau_D> \) | \( F_{xxx} \) |
|----------------|--------------------------|--------------------------|-------------|--------------|------------|
| 0201/0502      | 3.1425                   | 2.8774                   | 0.9791      | 2571.57      | 0.52       |
| 0302/0602      | 5.0585                   | 0.9528                   | 0.9806      | 2572.21      | 0.84       |
| 0302/0301      | 4.6532                   | 1.3574                   | 0.9807      | 2572.35      | 0.77       |
| 0302/0501      | 4.1317                   | 1.8785                   | 0.9808      | 2572.43      | 0.69       |

doi:10.1371/journal.pone.0107310.t003

doi:10.1371/journal.pone.0107310.t002
The next step consisted in simultaneously exploiting the \( \tau_{D,\text{Probes}} \) data pertaining to both probes. With this aim, we devised the graph shown in Fig. 5, in which each genotype is represented by a point whose abscissa is given by the \( \tau_{D,\text{Probe2}} \) value and whose ordinate is \( \tau_{D,\text{Probe1}} \). Moreover, for any heterozygous genotype DQB1-0XXX/DQB1-0YYY, the fraction \( F_{\text{XXX,\text{Probe1}}} \) of Probe1 hybridized to the DQB1-0XXX allele can be estimated graphically from Fig. 4. The relevant segments are indicated in Fig. 4 for the same DQB1-0201/DQB1-0502 sample also used to exemplify the constrained double-exponential fit procedure. Namely, \( F_{\text{XXX,\text{Probe1}}} \) (\( F_{\text{201,\text{Probe1}}} \) in the example) is found by dividing the difference (thick vertical segment in Fig. 4) of the ordinates \( Y_{\text{XXX,YYY}} \) and \( Y_{\text{YYY,YYY}} \) of the points corresponding to the heterozygous genotype (\( Y_{\text{201,502}} \) in the case under discussion) and to the homozygous DQB1-0YYY/DQB1-0YYY genotype (\( Y_{\text{502,502}} \) in the example) by the difference (thin vertical segment).

### Table 4. Initial amplitudes and \( \chi^2 \) values yielded by the fitting of the Probe2 data for the genotypes surviving the constrained fit analysis with both probes.

| Genotype       | \( f_{\text{xxx}} \times 10^4 \) | \( f_{\text{YYY}} \times 10^4 \) | \( \chi^2 \) | \( <\tau_D> \) | \( F_{\text{xxx}} \) |
|----------------|-------------------------------|-------------------------------|--------------|----------------|----------------|
| 0502/0201      | 5.3401                        | 4.6075                        | 0.9789       | 2220.76        | 0.54           |
| 0301/0302      | 5.8767                        | 4.0303                        | 1.0057       | 2190.09        | 0.59           |
| 0602/0302      | 6.9683                        | 2.9263                        | 1.0094       | 2281.25        | 0.70           |
| 0501/0302      | 7.1325                        | 2.7607                        | 1.0099       | 2281.37        | 0.72           |

In the last two columns the calculated values of the average lifetime and fraction of probe hybridized to the DQB1-0XXX allele are reported.

### Table 5. \( <\tau_D> \) values resulting from single-exponential fit of the decay data obtained for Probe1 (second column) and Probe2 (fourth column) made to react with the “heterozygous” oligonucleotide mixtures.

| Genotype       | \( <\tau_D,\text{Probe1}> \) | \( \sigma \) | \( <\tau_D,\text{Probe2}> \) | \( \Sigma \) |
|----------------|-------------------------------|--------------|-------------------------------|--------------|
| 0201/0501      | 2608                          | 7            | 2376                          | 12           |
| 0201/0302      | 2662                          | 7            | 2533                          | 11           |
| 0201/0502      | 2566                          | 9            | 2279                          | 11           |
| 0201/0503      | 2562                          | 10           | 2474                          | 8            |
| 0201/0602      | 2570                          | 10           | 2379                          | 9            |
| 0201/0301      | 2594                          | 5            | 2344                          | 5            |
| 0201/0402      | 2593                          | 11           | 2247                          | 10           |
| 0302/0402      | 2527                          | 9            | 2223                          | 6            |
| 0302/0501      | 2552                          | 8            | 2335                          | 7            |
| 0302/0301      | 2536                          | 9            | 2308                          | 4            |
| 0302/0502      | 2499                          | 11           | 2264                          | 7            |
| 0302/0503      | 2508                          | 10           | 2440                          | 4            |
| 0302/0602      | 2511                          | 11           | 2334                          | 5            |
| 0501/0502      | 2455                          | 8            | 2111                          | 7            |
| 0501/0503      | 2470                          | 9            | 2288                          | 10           |
| 0501/0602      | 2464                          | 12           | 2196                          | 9            |
| 0501/0301      | 2494                          | 9            | 2155                          | 10           |
| 0501/0402      | 2471                          | 6            | 2075                          | 10           |
| 0502/0602      | 2413                          | 11           | 2099                          | 6            |
| 0502/0503      | 2401                          | 9            | 2203                          | 6            |
| 0502/0301      | 2445                          | 11           | 2076                          | 6            |
| 0502/0402      | 2423                          | 12           | 1995                          | 4            |
| 0503/0301      | 2453                          | 6            | 2252                          | 5            |
| 0503/0602      | 2425                          | 3            | 2280                          | 8            |
| 0503/0402      | 2445                          | 6            | 2169                          | 6            |
| 0602/0402      | 2431                          | 11           | 2062                          | 8            |
| 0602/0301      | 2447                          | 9            | 2128                          | 6            |
| 0301/0402      | 2470                          | 3            | 2044                          | 7            |
between the ordinate $Y_{XXX,XXX}$ of the point representing the homozygous genotypes DQB1-0XXX/DQB1-0XXX ($Y_{201,201}$) and $Y_{YYYY,YYY}$. The same relation holds between the fraction $F_{XXX,\text{Probe2}}$ of Probe2 hybridized to DQB1-0XXX and the abscissa of the above-mentioned points of Fig. 4. Notably, for all the 10 heterozygous samples which were submitted to constrained double exponential fit analysis, the so derived values of $F_{XXX,\text{Probe1}}$ and $F_{XXX,\text{Probe2}}$ (in the example $F_{201,\text{Probe1}} = 0.50, F_{201,\text{Probe2}} = 0.52$) roughly coincide with those extrapolated from the constrained double exponential fitting with $\tau_{D,XXX}, \tau_{D,YYY}$ values proper of the correct genotype, but are very different from those yielded by all the other constrained fits fulfilling the $f_{XXX}$ and $f_{YYY}$ self-consistency criteria (see last columns of Tables 4 and 5 for the DQB1-0201/DQB1-0502 sample).

In Fig. 5, the rectangles surrounding the points indicate the confidence intervals ($\pm \sigma$) in the determination of $\langle \tau_D \rangle$.
Thus, superposition of two or more rectangles would denote ambiguous tagging of the corresponding genotypes. As it is apparent from Fig. 5, this situation holds for no genotype. This demonstrates our ability of recognizing any possible genotypic variant of the DQB1 trait encompassing codons 52–57 through combined analysis of the TCSPC data obtained with Probe1 and Probe2 within a confidence of one standard deviation. It must be said that such a degree of confidence is rather poor for a clinical test, corresponding to a probability to measure a τD value outside the rectangles >0.33. However, although not decisive, the above discussed method might be exploited as a preliminary screening tool, by virtue of its applicability on negligible amounts of untreated cell lysate, which constitutes a unique prerequisite for achieving superior readiness and cost-effectiveness with respect to traditional, PCR-based genotyping techniques. Moreover, if we limit our analysis to selected genotypes of cardinal clinical relevance, discrimination is feasible to a much higher degree of confidence. Namely, the genotypes DQB1-0201/DQB1-0201, DQB1-0201/DQB1-0302, and DQB1-0402/DQB1-0402, which are correlated with the maximum risk factors for celiac disease and IDDM, for Graves disease, and for biliary cirrhosis, respectively, are unambiguously typed even if the required confidence interval is extended to ±3σ, corresponding to a probability to measure a <τD> value outside the rectangles <0.003. Furthermore, with the same degree of confidence the ambiguity in recognition of DQB1-0602/DQB1-0602 genotype, encoding enhanced proneness to multiple sclerosis, sarcoidosis, and narcolepsy, involves three other possible genotypes only, namely DQB1-0502/DQB1-0503, DQB1-0602/DQB1-0501, and DQB1-0503/DQB1-0402. Finally, the clustering properties inherited by the one-dimensional <τD, Probe1> topology are conserved also by the distance distributions given by the two-dimensional Euclidean metrics applied on the (<τD, Probe1>; <τD, Probe2>) vectors. As a consequence, the genotypes originally comprehended within the green cluster, and predisposing to Graves and celiac disease, remain recognizable within the confidence interval of ±2σ, with the only ambiguity between the pathogenically relevant genotype DQB1-0201/DQB1-0502 and the non-predisposing genotype DQB1-0302/DQB1-0301. Obviously, removal of such degeneracy by submission of the patients suspected to bear the former genotype to further genetic screening with more refined techniques would be much easier than ab-initio traditional genotyping. The same can be said for the blue cluster, for which extension of the confidence interval to ±2σ only introduces ambiguity in the recognition of the pathogenically relevant genotypes DQB1-0302/DQB1-0402 and DQB1-0302/DQB1-0502, whose rectangles partially superimpose to those of the non-predisposing genotypes DQB1-0501/DQB1-0501 and DQB1-0501/DQB1-0503, respectively.

Conclusions

We designed two oligonucleotide probes specific for the highly polymorphic trait of the DQB1 gene encompassing codons 52–57, dual-labelled at their opposite ends with a fluorescence resonance energy transfer donor-acceptor pair. A full characterization of the fluorescence response of the donor when the probes are made to hybridize to oligonucleotide mixtures mimicking all the possible genotypes of that trait (8 homozygous and 28 heterozygous) was undertaken. We demonstrated that each genotype can be effectively tagged by the fluorescence decay constants extrapolated from the time-correlated single-photon counting data obtained with both probes. The probability of misassignments is >5% (corresponding to a ΔτD/2σ value <2) for 18 out of 36 genotypes.

Figure 5. Cross combination of the <τD> values obtained with the two probes allows discrimination of all possible DQB1 genotypes. Here, each genotype is represented by a point whose abscissa is given by the <τD, Probe2> value and whose ordinate is <τD, Probe1>. The half-base and half-height of the rectangles surrounding each point represent σ(<τD, Probe2>) and σ(<τD, Probe1>), respectively. Discrimination is warranted by lack of superposition among different rectangles.

doi:10.1371/journal.pone.0107310.g005
We are aware that such a poor degree of reliability in the attribution of genotypes makes the method unfit to serve as a self-sufficient diagnostic tool in clinical practice. Nonetheless, the readiness and cost-effectiveness of the method, combined with its provenance to be performed on untreated cell lysates, make it a valuable alternative to current genotyping techniques for preliminary wide-range population screening for genetic susceptibility to autoimmune diseases as severe as multiple sclerosis, celiac disease, narcolepsy, diabetes and several others.

Supporting Information

File S1 Details on constrained double-exponential fit procedure applied to DQB1-0201/DQB1-0502.

Genotyping of a polymorphic human gene trait by time-resolved FRET

Acknowledgments

We sincerely manifest our gratitude to Prof. Roberto S. Accolla and Dr. Giovanna Tosi (University of Insubria) for introducing us, years ago, to the DQB1 polymorphisms topic. The fruitful discussions and critical review of the manuscript by Prof. Francesco Mantegazza (University of Milan Bicocca) is also acknowledged. Thanks go to Dr. Angelo Masperso (University of Insubria) for making the Thermonucl fingerprinting method available for samples annealing.

Author Contributions

Conceived and designed the experiments: LN MB. Performed the experiments: ET ML LN. Analyzed the data: ET ML LN DS VC. Contributed reagents/materials/analysis tools: MB. Contributed to the writing of the manuscript: LN MB VC DS. Wrote Matlab routine for data analysis: ET ML. Suggested and implemented cluster analysis: ET.

References

1. Sachidanandam R, Weissman S, Schmidt EK, Kukar JL, Stein LD, et al. (2001) A map of human genome sequence variation containing 1.4 million single nucleotide polymorphisms. Nature 409: 928–933.
2. Schork NJ, Fallin D, Lynchburg S (2000) Single-nucleotide polymorphisms and the future of genetic epidemiology. Clin Genet 58: 250–264.
3. Greenblatt MS, Bennett WP, Holstein MM, Harris EC (1994) 53 Cancer Res 54: 4855–4878.
4. Marsh S, Kwock P, McLeod HL (2002) SNP databases and pharmacogenetics: great start, but a long way to go. Hum Mutat 19: 170–179.
5. German R (1994) MHC-dependent antigen processing and peptide presentation: providing ligands for T lymphocyte activation. Cell 76: 297–299.
6. Arranz E, Telleria JL, Sanz A, Martin JF, Alonso M, et al. (1997) HLA-DQA1*0501 and DQB1*02 homozygosity and disease susceptibility in Spanish celiac patients. Clin Immunogenet 14: 216–219.
7. Pedrano LM, Dema B, Lopez-Larios A, Malduna C, Bodas A, et al. (2012) HLA and celiac disease susceptibility: new genetic factors bring open questions about the HLA influence and gene-dosage effects. PLoS ONE 7: e41043.
8. Mignini F, Pizzi A (2012) HLA-DQA1 and HLA-DQB1 in celiac disease predisposition: practical implications of the HLA molecular typing. J Biomed Sci 19: 80.
9. Mantovani V, Corazza GR, Bragliai M, Frisoni M, Zaniboni MG, et al. (1993) Asp57-negative HLA DQB chain and DQA1*0501 allele are essential for the onset of DQw2-positive and DQw2-negative celiac disease. Clin Exp Immunol 91: 153–156.
10. Kaushansky N, Altman DM, David CS, Lassmann H, Ben-Nun A (2012) DQB1*0602 rather than DRB1*1501 confers susceptibility to multiple sclerosis-like disease induced by protocollpid protein (PLP). J Neuroinflammation 9: 29.
11. Kaushansky N, Altman DM, David CS, Lassmann H, Ben-Nun A (2009) HLA-DQA1*0501 determines disease susceptibility in a new “humanized” multiple sclerosis model in HLA-DR15 (DRB1*1501; DQB1*0602) transgenic mice. J Immunol 183: 3531–3541.
12. Santamaría P, Barboza JJ, Linndröm AL, Lemke TA, Goetz FC, et al. (1994) HLA-DQBI-associated susceptibility that distinguishes Hashimoto’s thyroiditis from Graves’ disease in type I diabetic patients. J Clin Endocrinol Metab 78: 478–483.
13. Pelín Z, Guiffreminault C, Risch N, Grumet FC, Mignot E (1998) HLA-DQBI*0602 homozygosity increases relative risk for narcolepsy but not disease severity in two ethnic groups. Tissue antigens 51: 96–100.
14. Olerup O, Edvard Smith CI, Bjorkander J, Hammarström L (1992) Shared HLA class II-associated genetic susceptibility and resistance, related to the HLA-DQBI gene, in IgA deficiency and common variable immunodeficiency. Proc Natl Acad Sci USA 89: 10633–10637.
15. Immunuzzi MC, Mallarik MJ, Poisson LM, Rybicki BA (2003) Sarcoidosis susceptibility and resistance HLA-DQB1 alleles in African Americans. Ann J Respir Crit Care Med 167: 1292–1299.
16. Underhill J, Donaldson P, Bray G, Dobert D, Portman B, et al. (1992) Susceptibility to primary biliary cirrhosis is associated with the HLA-DRB1*0402 haplotype. Hepatology 16: 1404–1408.
17. Vázquez-García MN, Alarcón C, Olivero A, Debez H, Pérez-Luque E, et al. (1998) MHC class II sequences of susceptibility and protection in Mexicans with autoimmune hepatitis. J Hepatol 29: 885–890.
18. Congia M, Clemente MG, Desi C, Cucca F, Mazzoleni AP, et al. (1996) HLA class II genes in chronic hepatitis C virus-infection and associated immunological disorders. Hepatology 24: 1338–1341.
19. Gavrilov DK, Karaveva TL, Dedov II, Sergeev AS, Nosikov VV (1994) Frequency analysis of HLA-DQA1 and HLA-DQB1 gene alleles and susceptibility to type 1 diabetes mellitus in Russian patients. Acta Diabetol 31: 62–69.
20. Erlich HA, Griffith RL, Bugawan TL, Ziegler R, Alper C, et al. (1991) Implication of specific DQB1 alleles in genetic susceptibility and resistance by identification of IDDM siblings with novel HLA-DQB1 allele and unusual DR2 and DR3 haplotypes. Diabetes 40: 478–481.
21. Kim S, Misra A (2007) SNP genotyping: technologies and biomedical applications. An Rev Biomed Eng 9: 289–320.
22. Tyagi S, Bratu DP, Kramer FR (1998) Multicolor molecular beacons for allele discrimination. Nat Biotechnol 16: 49–53.
23. Gaylord BS, Massie MR, Feinstein SC, Bazan GC (2003) SNP detection using peptide nucleic acid probes and conjugated polymers: applications in neurodegenerative disease identification. Proc Natl Acad Sci USA 100: 34–39.
24. Al Attar HA, Norden J, O’Brien S, Monkman AP (2008) Improved single nucleotide polymorphisms detection using conjugated polymer/surfactant system and peptide nucleic acid. Biosensors & Bioelectronics 23: 1466–1472.
25. Andreoni A, Bondani M, Nardo L (2009) Feasibility of single nucleotide polymorphism genotyping with a single probe by time-resolved Forster resonance energy transfer. Mol Cell Probes 23: 119–121.
26. Andreoni A, Bondani M, Nardo L (2009) Time-resolved FRET method for typing polymorphic alleles of the human leukocyte antigen system by using a single DNA probe. Photochem Photobiol Sci 8: 1202–1206.
27. Nardo L, Tosi G, Bondani M, Accolla RS, Andreoni A (2012) Typing of a polymorphic human gene conferring susceptibility to insulin-dependent diabetes mellitus by picosecond-resolved FRET on non-purified/non-amplified genomic DNA. DNA Res 19: 347–355.
28. Nardo L, CAMERA N, TOET E, Bondani M, ACCOLLA RS, TOSI G (2013) Time-resolved Forster resonance energy transfer analysis of single-nucleotide polymorphisms: towards molecular typing of genes on non-purified and non-PCR-amplified DNA. J Mol Biol Res 3: 15–22.
29. Andreoni A, Bondani M, Lamperti M, Nardo L (2014) Typing of polymorphic human genes by picosecond-resolved Forster energy transfer. IEEE J Sel Topics Quant Elec 20: 100–107.