Title

IntelliEppi: Intelligent reaction monitoring and holistic data management system for the molecular biology lab

Authors
Arthur Neuberger[1, +], Zeeshan Ahmed[2, +], Thomas Dandekar[3, *]

Affiliations
1. Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1PD, UK.
2. School of Medicine, University of Connecticut Health Center, 195 Farmington Ave, Farmington, 06032, CT, USA.
3. Department of Bioinformatics, Biocenter, University of Wuerzburg, 97074, Wuerzburg, Germany.

[+]equally contributing first authors
[*]Corresponding author(s): Thomas Dandekar (dandekar@biozentrum.uni-wuerzburg.de)

Abstract
Daily alterations of routines and protocols create high, yet so far unmet demands for intelligent reaction monitoring, quality control and data management in molecular biology laboratories. To meet such needs, the "internet of things" is implemented here. We propose an approach which combines direct tracking of lab tubes, reactions and racks with a comprehensive data management system. Reagent tubes in this system are tagged with 2D data matrices or imprinted RFID-chips using a unique identification number. For each tube, individual content and all relevant information based on conducted experimental procedures are stored in an experimental data management system. This information is managed automatically but allow scientists to engage and interfere via user-friendly graphical interface. Tagged tubes are used in connection with a detectable RFID-tagged rack. We show that reaction protocols, HTS storage and complex reactions are easily planned and controlled.

Introduction

The Internet of things is increasingly becoming a tangible reality of the 21st century with examples from the modern factories for automatic assembly lines that work complementarily to a supportive smart-tag based storage system (1). It is not restricted to certain technological areas, instead, with the addition of some creativity, it can be applied to almost every complex process in which electronic devices are involved (2). It is therefore rather surprising that such a development is still missing in most of the modern biological laboratories, where automation is becoming more and more visible through the innovation, improvement of kits, state-of-the-art omics technology and laboratory information management systems (LIMS).

First attempts to create integrated LIMS can be traced back to the early 80's in the form of early patents, whereas, real marketing efforts only became noticeable in recent years, when some leading laboratory equipment suppliers started to develop technology-based intelligent laboratory systems. A good example for this development is the advent of fully automatic systems like Eppendorf's robotic workstation "epMotion" (3) or Dornier-LTF GmbH's pipetting robot "PIRO" (4). An automated process for analysis has been proposed in several patents. For instance, Lang and colleagues (5), included "storage" of experimental lab data in the memory of a contactless chip card or a barcode attached to the corresponding reaction vessel. Moreover, RFID (Radio-Frequency Identification) technology has already

Neuberger, A., Ahmed, Z., and Dandekar, T.

52 successfully been implemented commercially in the electronic identification and
53 administration of blood donations in transfusion medicine (6, 7).
54 Even though lab machinery and devices work quite efficiently within their own technological
55 boundaries but almost no integrative interconnection between these smart tools has been
56 established to date. These "closed" systems are unfortunately not yet smart and integrated
57 enough to fully replace the operator and his manual interventions in an environment like the
58 academic molecular biological lab. This demand for facilitated and smart, yet partly manually
59 controlled half-automation of experimental processes could only be met in a semi-open
60 system. Most of the available sample data management systems lack full integrity, i.e. a
61 smart solution that would integrate common devices used on a daily basis, like pipettes and
62 test tubes, into a semi-automatic system.
63 Facing this stagnation of integrative innovation in the laboratory equipment market, we
64 propose a holistic system that we call "IntelliEppi" (8). It brings the concept of internet of
65 things to the labs. The system is a proof of concept for an internet of things lab
66 management, which, combines economic and easy tagging via two-dimensional data matrix
67 codes or printable RFID tags, with a smart reagent tube logistics and experimental data
68 management software. It supports tracking of all components can be tracked with the help
69 of IntelliEppi, enhance to deliver reagents on time, at the right place, allowing complex
70 synthesis processes and improved quality in the process and its controls.
71 IntelliEppi allows to store and modify molecules by monitoring, guiding and storing reaction
72 vessels in defined positions, using a reader-system (via a matrix printer or RFIDs) and the
73 power of a versatile program code. It leads to better quality, reproducibility, and
74 transparency in biochemical and molecular biology experiments (9). It incorporates product
75 data management (PDM) from the start and allows reaction tube lifecycle management
76 including long term storage processes, tracking of resources and tubes, and a scheduler.
77 Using IntelliEppi simple reactions and really complex omics experiments can be planned and
78 different tagging strategies in the lab can broadly be explored. All this is made available with
79 the hope to promote the vision of an internet of things from the test tube to the laboratory
80 scale, bundled here in the concept of the intelligent reaction vessel monitoring and actually
81 doing all reaction steps in various experiments. We tested all ingredients and make them
82 available, which includes software, different tagging strategies, performance data, tutorials,
83 labels as well as pseudocode and examples on the synthesis processes.
84
85 ## Results
86
87 **IntelliEppi components and usage:** The system IntelliEppi is holistic and comprehensive
88 to provide an internet of things for the molecular biology laboratory and the whole life cycle
89 of the reagent tube:
90 1. Reagent tubes in this system are tagged with cheap, yet resistant printable plastic
91   RFID chips or 2D data matrices that are marked (e.g. using a laser) onto the top of
92   their lid. These tags assign and store a unique identification number through which
93   each individual tube's content and all relevant information on conducted
94   experimental procedures of the sample inside the tube can be requested on demand
95   and instantly gathered (Fig. 1(A-C)).
96 2. All information is stored in a database that is managed semi-manually by the
97   scientist. A user-friendly GUI allows constant information gathering and data editing.
98 3. Tagged tubes are stored in a smart tube rack which can be tracked via its own RFID-
99   tag. This happens both physically via long-distance scanning and digitally with all
100   relevant information connected to the tag displayed when a specific sample tube
101   needs to be found and identified.

2

102   4. The rack can be scanned for information on tube contents and the place of the tube
103   inside the rack. Alpha-numerical- or colour-coding on the rack assign each tube an
104   individual slot on the rack.
105   5. The rack itself, inside the fridge for instance, can be easily identified via a lighting up
106   LED, which is activated when it is scanned. This way, no samples are lost due to
107   unknown location or erased labels. This makes tracking of old samples and
108   connected information easier and faster.
109   6. The database stores all relevant information for individual tubes, entire racks, and
110   chemicals (buffers, reagents etc.). All components are stored in tagged tubes or
111   flasks and can therefore be integrated via their identifiers into an interactive digital
112   experimental protocol. This allows both precise design of an experiment and
113   guidance through it, with possible alterations by the user when conducting the
114   experiment. Here, performed steps and other relevant information are updated to
115   the database and stored for every individual tube.
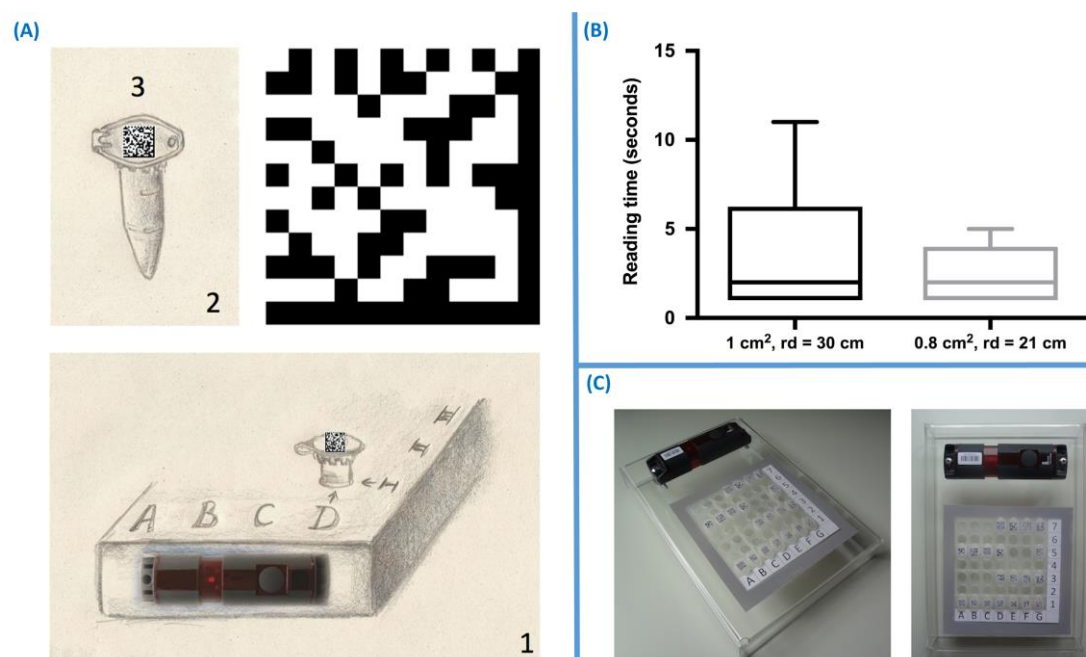
116

117   **Figure 1. (A)** Tagged SMARTtube, (B) box plot of scanning time data matircs, and (C)
118   SMARTrack prototype.

119   **IntelliEppi workflow and validation:** IntelliEppi's power is demonstrated with the aid of
120   conceptual coding examples for different chemical reactions. The combination of software,
121   matrix code or RFID guiding and intelligent tubes achieves an equivalent of NCR guided
122   technology for biochemical reactions.
123   A standardised process flow is schematically outlined in Fig. 2(A), individual steps are shown
124   in Fig. 2(B). As an appropriate practical example for a complex synthesis we have chosen the
125   activity and kinetics real-time monitoring of the T4 polynucleotide kinase reaction. It is based
126   on a singly labelled DNA-hairpin smart probe coupled with λ exonuclease cleavage (12) (our
127   demonstrator model system). In their approach, Song and Zhao designed a smart probe
128   (labelled oligonucleotides with a hairpin shape) with a fluorophore at the 3'-phosphate end.
129   The Fluorescence is quenched by a guanine-triplet at the terminal 5'-hydroxyl group (12). In
130   the presence of ATP, this 5'-hydroxyl group of the smart probe is then phosphorylated by the
131   T4 polynucleotide kinase (Fig. 2(C)). In a second step, fluorescence enhancement is caused
132   when the resulting 5'-phosphoryl end is cleaved by the λ exonuclease. The latter is a $5 \rightarrow 3$
133   exonuclease with a preference for phosphate moieties at the 5' of double-stranded DNA

Neuberger, A., Ahmed, Z., and Dandekar, T.

134 ends that degrades these double-stranded DNA while yielding mono-nucleotides and single-
135 stranded DNA. The detection process is carried out in a real-time PCR instrument as this
136 offers smooth temperature control, sealed tubes, and high-throughput detection (12).
137 This experiment can be divided into various steps (Fig. 2(B)), each of which is associated with
138 important information. In the first stage, a test tube is tagged with an individualized 2D
139 barcode using IntelliEppi's "Generator" function. When using the latter, the user defines the
140 experiment, e.g. "T4 polynucleotide kinase (pnk) reaction", and manually fills out the desired
141 range of accompanying options, such as the conducting researcher's name, the SmartTube's
142 No. as well as the SmartRack's No. and SMARTtube storage position within this, and the date
143 and time of the experiment. By clicking on the 2D barcode symbol, the user generates a 2D
144 data matrix which stores the information from all fields in a compact, yet comprehensible
145 code format like "ST73|SR02|A5|T4 pnk reaction|19/07/2016 19:30:15". The latter, being
146 encoded in the actual 2D tag label, can later be easily read for identification of the
147 SMARTtube, e.g. in a rack inside a fridge. A further click on the database symbol stores the
148 entire information package into the ExKPDM database for long term storage, identification
149 and data management in the future.
150 Fig. 1(A) illustrates the SMART reaction control. Template DNA and a nucleotide mix are then
151 added to the tagged tube (Position 1). If preferred by the user, this step and the following
152 ones are manually noted down in the corresponding ExKPDM file that was created for the
153 registered tube: In position 2, the T4 polynocleotide kinase is added and therewith the
154 reaction started. The reaction is stopped after 30 minutes (Position 3). The now labelled
155 reaction can then be added to the hybridization platform and maintained there for a total
156 reaction time of 24 hours (Position 4) before the filter is taken out (Position 5) and
157 transferred to the detector for read out (Position 6).
158 IntelliEppi is able to perform the described experiment whilst producing an up-to-date and
159 lasting live documentation. A single document file is semi-automatically updated at every
160 step in the reaction flow. After the experiment, the tube (containing the product of the T4
161 pnk reaction) can be tracked via its 2D tag using a scanner connected to the ExKPDM system
162 (run on a computer; scanner connected via USB boost or Bluetooth; more details on tagging
163 and tracking suppl. Material, part 1). The T4 polynucleotide kinase reaction SMARTtube can
164 be stored on a SMARTrack which is also registered in the system under the same experiment.
165 However, in this case, the SMARTrack's RFID tag is scanned by an RFID reader. The latter is
166 also connected to the ExKPDM system. The reader is, as it is the case for the optical 2D data
167 matrix scanner, connected via USB boost or Bluetooth with a computer. Identec Solutions'
168 system works with TCP/IP or COM connections. A reading device is currently under
169 development that can be directly connected to a computer (via USB boost).
170 ATP, the smart probe, T4 polynucleotide kinase, λ exonuclease, and other reagents used in
171 various experiments in this lab can be stored together in designated SMARTrack. The
172 following features are conceptual and under current implementation into the software
173 package: Reagents, buffers, and other chemicals could also be registered in the ExKPDM
174 chemical database. All chemicals (reagents, buffers etc.) would have their own individual
175 ID/EPC. Hence, when a new experiment is designed, using the experimental protocol
176 function of the ExKPDM, all experimental components (buffers, reagents etc.) as well as the
177 SMARTtubes, which these experiments are about to be performed in, would be manually
178 identified using a search query and then registered into the new experimental protocol. In
179 this protocol, all experimental steps, e.g. how much ATP shall be put in at which time point
180 into which SMARTtube, is clearly pre-defined before any action is taken. This protocol could
181 also be manually adjusted at every point of the experiment (as long as the user has the rights
182 to do this). In case of experiments in which certain timing points must be observed, a timing
183 function enables the user to plan this dimension of his experiments as well.

4
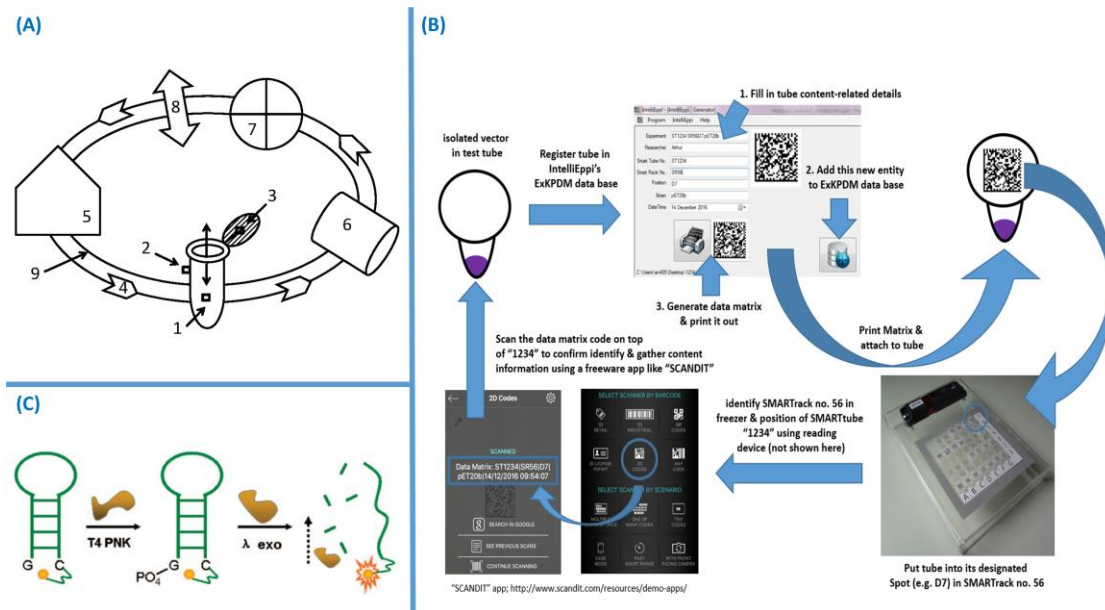
Figure 2. (A) The reaction cycle, (B) summary for one example of a simple application routine for the IntelliEppi system, and (C) T4 Reaction at the terminal 5'-hydroxyl group.

As previously mentioned, such a digital protocol would be interactive, changeable at any point and more like a programme that is running in parallel until it reaches its defined end. This programme would automatically notify the user when it is time for him to execute a timing-dependent step in the experiment. It would also provide information on the next step to execute. In the era of smartphones and tablet PCs, one might even consider a complementary app that informs the user everywhere and at any time. This way, experiments can be easily managed from abroad, or outside the lab, by an instructor or group leader. Every change of a SMARTtube's content is written into the experimental protocol and subsequently into the SMARTtube database. Next time when this SMARTtube is scanned, the user receives the updated information as well as the full experimental history of this tube (i.e. all executed steps, content changes together with the corresponding time points and maybe also the name of the person who conducted this step).

The internet of things for the laboratory requires integration of software and lab components. Thus, besides tagging and tracking for programming of reaction courses and integration of all components, the software is critical. SMARTtubes belonging to the same experiment or analysis fraction are always stored in the same corresponding SMARTrack, except during their handling, i.e. when the content of a tube is changed or analysed for instance. The IntelliEppi laboratory database, as one of several software components of the ExKPDM system, stores the EPCs/IDs of each SMARTtube and –rack. The user who is trying to find an individual tube for further experimental procedures could either directly search for the tube by typing the ID key or EPC into a search tool, or, alternatively, search for the corresponding SMARTrack (in which this tube was stored) using the SMARTrack-ID. As an output of such enquiry for an individual SMARTtube, or a whole SMARTrack, an interactive window opens that displays all information concerning this SMARTtube or, in the case of a search for a whole SMARTrack, all registered SMARTtubes belonging to this rack, respectively. In the latter case, manual picking of an individual SMARTtube (via SMARTrack window) would display the lab history of this tube (as it is also the case for the output of a manual search for this SMARTtube).

Fig. 1(C) illustrates a hypothetical mobile version of IntelliEppi. Information was analysed on the position of the searched tube in this SMARTrack (row and column, e.g. "D1"); the place where (in which fridge, floor etc.) the SMARTrack was located (e.g. "Fridge II, Lab. 3, 2nd

5

218 Floor"); the time when the sample had been created (e.g. "Experiment started: 17:10:00
219 04.07.2013"). The system considers also the experiment this SMARTtube belongs to (e.g. "T4
220 pnk activity"), when and how the content was changed. All this is analysed via an interactive
221 experimental history sliding window (see "Add 0.5 mL Buffer 17:18:23 04.07.2013") using a
222 user-friendly GUI. Cross-links lead to additional information: e.g., on the experimental
223 protocol (created by the user using protocol function of ExKPDM software), on chemical
224 knowledge of the reagents used for analysis (see stylised selector: "Go to chemical
225 knowledge base") or content change/protocol following, on EPCs/IDs and technical
226 information of lab equipment used in connection with this tube or its content respectively.
227 A tube´s assigned location in a rack can be easily and economically controlled using colour-
228 and/or numerically coded rack surfaces. For example, an individual tube could be allocated a
229 pre-defined storage position like D1 or red/D or red/1 in the SMARTrack. Columns and rows
230 are thus labelled and/or colour-coded in the rack. Obviously, a user looking at a SMARTrack
231 filled with SMARTtubes (datamatrix coded for example) cannot possibly distinguish between
232 the tubes on the basis of different codes on the top of lids. Thus, it is required that every
233 individual tube, after being taken out of the rack, e.g. for centrifugation or pipetting, is
234 always put back in its pre-defined position in the rack, like D1. Information on which the right
235 place for an individual tube is, could be instantly recalled by scanning its tag.
236
237 **Designing complex reactions with IntelliEppi.** More complex reactions often involve
238 molecular biology genetic engineering. As an example (Fig. 3 (A)), a fusion construct consists
239 of PCR generated fragment with a light-gated BLUF domain is attached to a T4
240 polynucleotide kinase. This makes the T4 polynucleotide kinase light-dependent and so light
241 can be used as an additional external control in the system. The final construct is stored in
242 position end.
243
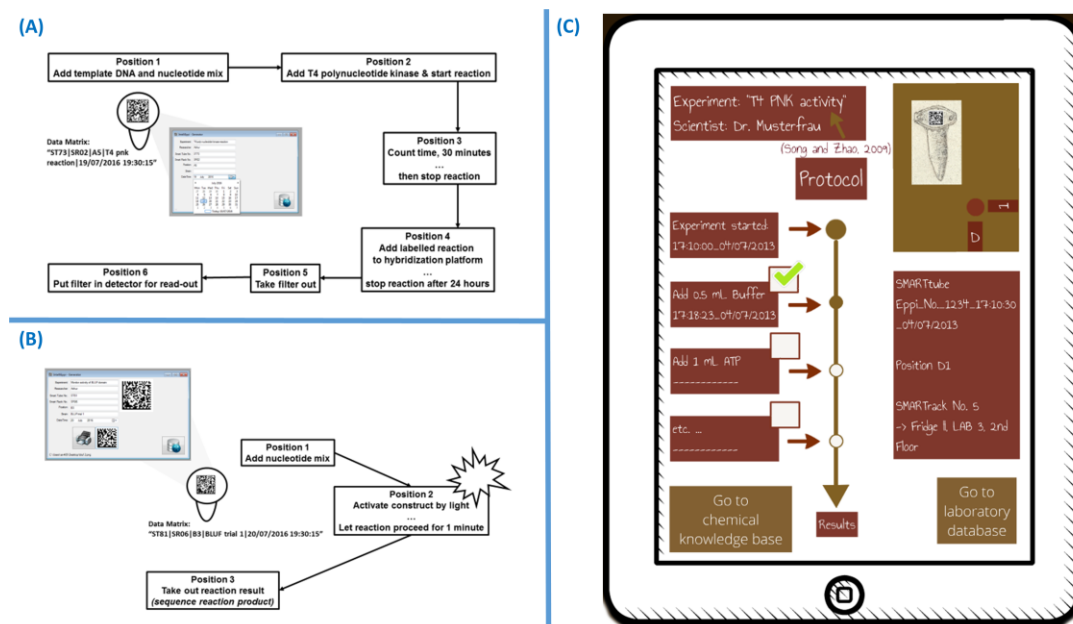
245 **Figure 3**. (A) The T4 polynucleotide kinase reaction is given as an example for chemical
246 reaction pathway engineering using IntelliEppi. (B) BLUF domain fusion activity testing as an
247 example for chemical reaction path engineering using IntelliEppi. (D) Abstract mobile
248 IntelliEppi-ExKPDM user interface.
249

250 This more complex program consists of the following subroutines: Cloning of BLUF domain
251 (Fig. 3(B)), 1: Template DNA extraction from *E. coli*, 2: Add suitable PCR primers; PCR the
252 BLUF domain, 3. Easy cloning step to get BLUF domain in vector of choice. 4: Template T4

6

253 DNA, PCR the T4 polynucleotide kinase, 5: Easy cloning step to get T4 kinase domain in
254 vector of choice, 6: Transform the vector into *E. coli*), and Expression and purification of the
255 BLUF domain (Fig. 3(B)), 7: Express the fusion construct, and 8: Purify the fusion construct
256 on a nickel column). Finally, the activity of the new construct is tested in an example where
257 the IntelliEppi helps to monitor the underlying chemical reaction pathway.
258 Another area where the additional smart control and the IntelliEppi system with software is
259 handy is for complex synthesis processes. Examples explored here are dendrimer synthesis, a
260 DNA macrame or RNA designer aptamers wired for a logical gate or array.
261
262 ## Discussion
263
264 Recent alternatives to IntelliEppi's software and monitoring components include the
265 pipetting robot Pirot as a bench worker robot and the Emerald cloud where you order a
266 service with a programming language that helps to manage different types of laboratory
267 work. Several further software languages allow to formulate experiments. For instance,
268 there is Biocoder, a programming language for standardizing and automating biology
269 protocols (11). The PaR-PaR laboratory automation platform (12) features the "biology-
270 friendly high-level robot programming language PaR-PaR (programming a Robot;
271 http://prpr.jbei.org/). Furthermore, there is the formalization language EXACT (13).
272 Moreover, you can order experiments from Emerald cloud even though there is no
273 automation for this. However, these are only partial solutions. We provide an integrated
274 solution: IntelliEppi complements the efforts described above by allowing the detailed design
275 and step-by-step monitoring of complex biochemical laboratory experiments and by using a
276 specific programming language focused on position and function codons for guiding and
277 monitoring the reaction path together with the reaction vessel (the "Eppendorf tube") in an
278 intelligent way. A less technologically advanced but also much cheaper and "semi-open"
279 approach for half-automated sample management is nowadays realized in the use of label
280 printers for individual samples (14). In some cases, these printers are sold with basic sample
281 data management software (e.g. Brady Laboratory Labels).
282 The identification of molecular biology laboratory samples in a container using RFID-
283 technology, has been proposed by Excoffier and colleagues (15). A combination of barcode
284 and RFID tagging on the same test tube has also been proposed in form of a system where
285 test tube identification data is encrypted in a barcode whereby additional data is stored on
286 the RFID tag (16). An advancement of the latter is a system comprised of a rack holding
287 element and, adjacent to this, of a carrier for an antenna for the wireless reading of an RFID
288 tag attached to a test tube (17). In a different, further automated system, test tubes, tagged
289 via barcode, are transported in a RFID-tagged conveying device via a conveyor belt to pre-
290 testing, testing, and post-testing stations. Corresponding devices are mounted to the system
291 whereby identification and control of action in this system are based on RFID tagging of the
292 conveying devices (18). With a view to relevant patents for RFID technologies, printable RFID
293 transponders are likely to constitute one of the most promising approaches (10). The
294 corresponding printing technology has already been proposed in a separate patent (19). This
295 is only an excerpt from the lively patent scene for RFID- and barcode-based life science
296 applications.
297 Some proposed approaches aim to combine tagging with simultaneous content analysis or
298 reaction monitoring, like those using encapsulated microprobes and lab-on-chip systems.
299 The latter particularly apply to diagnostics. A system-on-chip digital pH meter, for instance,
300 has already been successfully enclosed in a diagnostic capsule for in vivo diagnostics of
301 gastro-intestinal conditions and diseases (20). An even broader spectrum of analytical
302 measurements is offered by a micro-lab system that is encapsulated in a biocompatible shell
303 (21). Digital cloud-based lab management systems (Laband.me LTD is a good example for
304 such a provider; see also https://www.laband.me/) offer easy-to-use data recording and

305 analyzing tools, based on apps, cloud-storage, and electronic notebooks, making the daily
306 management, simple statistical analysis, or straightforward visualization of data easier.
307 However, none of the mentioned systems integrate system tagging, detection and analysis
308 into an intelligent and user-friendly management software solution.
309 Our software has the advantage to be easily, generically adapted to new, different
310 biochemical experiments and is easy to be scaled up. Given examples include the T4
311 polynucleotide kinase reaction cycle, a more complex genetic engineering example with PCR
312 protocol, and some complex synthesis examples. Furthermore, the tool also enables a library
313 generation for RNAseq, or omics-seq experiments, ClipSeq, ChIPseq, RIPseq. These protocols
314 are not shown here but are easily accomplished through this strategy, which affords relative
315 flexibility to the reaction tube position, typical reaction steps as well as storage and different
316 addition modification and filtering steps.
317 IntelliEppi has to be seen as a complementary part of a general effort to automate, monitor,
318 and design lab experiments. This includes the Emerald Labs where you can order
319 experiments from these. There is also the use of a robot scientist, well-known from an early
320 article (22), (23). The aim is again to free the mind for intelligent work. There is also the idea
321 to increase the flexibility in experimental design (as evidenced by our flexible code) pursued
322 by current efforts, for instance in instructing 3D printers (24), flexible synthesis-like software
323 (25) and HTS advances such as digital picoliter PCR(26), (27).
324 This is a first step in the direction towards a micro-factory environment. Further extensions
325 include the synthesis of even more complex compounds and further miniaturization.
326
327 ## Methods
328
329 **SmartTube and SMARTrack:** Test tubes (e.g. provided by the Eppendorf) are involved in
330 highly complex processes conducted in a modern biochemical or molecular biological lab. It
331 is therefore evident that these tubes act as the cardinal element in a lab logistics system.
332 Most laboratories are still required to manage their sample and data logistics by hand. Thus
333 far, the most innovative sample data management system to mention is based on label
334 printers that are capable of creating labels of alphanumerical content as well as 2D codes
335 which have a sticky back via which these can be attached to tubes. In a given situation, as
336 shown in Fig. 2(A); a tagged reagent tube's ("SmartTube") lid surface was tagged with a data
337 matrix. SMARTtube (sitting inside a SMARTrack) tagged with a data matrix ECC encoding
338 "ASCII" (ECC 200 for instance). SMARTrack, tagged via i-Q350L FLSensorSMART, with a
339 SMARTtube inserted at a designated position. SMARTtube tagged with ECC200-datamatrix.
340 ECC-datamatrix encoding "ASCII".
341
342 **Datamatrix coding:** An alphanumerical code like "ST1234SR56D7 17:10:38 03.09.13" for
343 instance can be encoded in a data matrix composed of 20 x 20 modules. This code can be
344 interpreted as: "SMARTtube no. 1234 in SMARTrack no. 56, positioned in row D and column
345 7 in this rack, last procedure was executed at 17:10:38 on 03.09.2013". A module size of 0.4
346 mm leads to a total code area size of 8.0 mm2. This dimension is small enough to be lasered
347 onto the top of a standard reagent tube lid. We tested the data matrix coding. It is large
348 enough to be read from the maximal distance of 30 cm (for a 1 cm2 matrix; 21 cm in case of
349 a 0.8 cm2 matrix) within 2 seconds on average using a simple freeware smartphone app (Fig.
350 2(B)).
351
352 **RFID-tagging:** In the case of a RFID tag printed on top of a tube lid, its electronic product
353 code stored in the memory of this RFID tag (that enables individualisation of every single
354 tube) could be read by an appropriate RFID reader device (depending on the operating
355 frequency in Hz). For this technology, we tested RFID tags made by Identec Solutions AG.
356 These are moderately sized RFIDs (6-10 cm) that are commercially used for location tracking

8

357 of lorries and are utilized in our system for reaction tube rack tracking in a laboratory setting.
358 In combination with a read-out device, these RFID tags were obtained from Identec Solutions
359 AG. Anticipating future developments, we also looked at the RFID potential from PolyIC's
360 imprintable RFID tags. Here, different polymer foils are coated with electronics – potentially
361 a very powerful technique that would enable submillimetre miniaturization of cheap tags for
362 reagent tubes. However, this technology is not yet in production mode. According to the
363 developing company, this technology will take another 3-5 years until it becomes available.
364 Hence, the best technology currently available features 3-4 mm RFIDs from Microsensys. As
365 mentioned earlier, 2D coding was used in our demonstrator model. Here, an individual tube
366 is allocated an individual ID key that is stored in an IntelliEppi database as a part of a whole
367 Experimental Knowledge and Product Data Management Software.
368
369 **Demonstrator Model:** For our demonstrator model, data matrices were generated using
370 the 2D barcode "Generator" module of the IntelliEppi software. These barcodes have the
371 following meta data: Experiment, Researcher, Smart Tube Number, Smart Rack Number,
372 Position, Strain and Data Time. SMARTtubes are always used in a logical connection with an
373 RFID-tagged rack, henceforth referred to as "SMARTrack". Here, Identec Solutions GmbH, for
374 instance, provided a battery-powered active tag, i-Q350L FLSensorSMART (Fig. 2(C)) for our
375 model system. This tag operates at 868 MHz (EU-compatible) with a localisation and a
376 read/write range response mode of several hundred meters in free air, a memory size of
377 10,000 bytes (user definable), a 48-bit fixed identification code, a replaceable Lithium
378 battery, and a special marker function with an operating frequency of 125 kHZ. Tag
379 dimensions are 137 x 37.5 x 26.5 mm and therefore suitable for placement at the short side
380 of common tube racks. Furthermore, an LED responds to tag scanning with a bright optical
381 signal when the tag is detected in our model. One can imagine that dozens of SMARTracks
382 are stored in a fridge at the same time. Scanning the content of this fridge for SMARTrack no.
383 56 for example, gives back an optical signal that allows quick identification of the right rack
384 (where the searched tube is stored). SMARTtubes and the SMARTrack, in which these
385 SMARTtubes are permanently stored, form an intelligent unit within the IntelliEppi system.
386 Together with a reader (and scanner) they form the hardware component of the IntelliEppi
387 system.
388
389 **Experimental Knowledge and Product Data Management Software:** This is henceforth
390 referred to as "ExKPDM", is the software component of this system. The ExKPDM software
391 components are shown in Fig. 3(A and B) as a conceptualized work-flow overview of basic
392 components, their cross-linking and interaction within IntelliEppi, and the resulting user
393 benefit. Labels are as follows: 1: RFID tag as a reusable probe insight the tube; 2: RFID tag
394 attached to or a 2D data matrix barcode printed/lasered on the surface of a tube; 3: RFID tag
395 incorporated into or a 2D data matrix barcode printed/lasered on the lid of a tube
396 (incorporation as part of manufacturing). 4: the process flow (not necessarily cyclic); 5:
397 storage module; 6: reaction module; 7: detection module; 8: "mini-factory", semi-open
398 system (user intervention is possible) for complex syntheses; 9: ExKPDM including IntelliEppi-
399 Tracking. (B) ATP + 5' − dephospho −DNA −→ ADP + 5' − phospho − DNA (Figure adapted
400 from (12)).
401
402 Specific software modules are developed according to user needs. The user feeds the
403 ExKPDM system with instructions or data via a user-friendly user interface. All further
404 processes run based on this interface in the background, not visible for the user as such, Fig.
405 3(C). Briefly, individual SMARTtubes are operated through the O-module. Data on each
406 SMARTtube is saved into the laboratory database D via the module for optimal SMARTtube
407 management. The chemical knowledge base enables the user, via the user interface to carry
408 out a system-integrated search in various chemical databanks. The data are also transferred

Neuberger, A., Ahmed, Z., and Dandekar, T.

409 into the laboratory database ingtegrating query results and programmed searches. Final
410 results of the search module appear via the user interface. Finite working cycle (life cycle)
411 data can be requested using 2D data matrix/RFID tagging of SMARTtubes and SMARTracks
412 linked to the ExKPDM.
413
414 **Code availability:** The executable of the ExKPDM system and a test database are available.
415 The source code will be made available in the same way upon acceptance of the manuscript
416 without any restrictions of usage. Tagging, protocols, RFID information, software and
417 tutorials are all available at: www.bioinfo.biozentrum.uni-
418 wuerzburg.de/computing/intellieppi
419 Furthermore, source code, setup and executable are all loaded up to the GitHub public
420 repository at https://github.com/drzeeshanahmed/IntelliEppi
421

Author contributions
423 A.N. did all work on the IntelliEppi Lab implementation including demonstrator, RFID tests,
424 data matrix tests and tagging. Z.A did all work on the software aspects of IntelliEppi including
425 the expert system and implementation of code. T.D. lead and guided the study, analyzed the
426 performance and data of the IntelliEppi system. All authors drafted the manuscript and
427 finalized it together.
428

References
430 1. Huang, G. Q., Zhang, Y. F. & Jiang, P. Y. RFID-based wireless manufacturing for walking-
431 worker assembly islands with fixed-position layouts. *Robot Cim-Int Manuf*. **23**, 469-77 (2007).
432 2. Atzori, L., Iera, A. & Morabito, G. The internet of things: A survey. *Computer networks*. **54**,
433 2787-805 (2010).
434 3. Spolaczyk, R. Apparatus for handling liquids and a process for operating the device.
435 https://www.google.com/patents/US6819437 (2004).
436 4. Becher, H. & Renz, M. Pipetting robot for laboratory use.
437 https://www.google.com/patents/USD674110 (2013).
438 5. Mann, K. H., Lang, A., Nowak, E., Sattler, S. & Schels, H. D. Verfahren zur Analyse von
439 Probenflüssigkeiten. https://encrypted.google.com/patents/EP0637750A2?cl=it (1995).
440 6. Davis, R., Geiger, B., Gutierrez, A., Heaser, J. & Veeramani, D. Tracking blood products in
441 blood centres using radio frequency identification: a comprehensive assessment. *Vox*
442 *sanguinis*. **97**, 50-60 (2009).
443 7. Hohberger, C., Davis, R., Briggs, L., Gutierrez, A. & Veeramani, D. Applying radio-frequency
444 identification (RFID) technology in transfusion medicine. *Biologicals: journal of the*
445 *International Association of Biological Standardization*. **40**, 209-13 (2012).
446 8. Neuberger, A., Ahmed, Z. & Dandekar, T. A system for non-contact monitoring of reaction
447 vessels with electronic bearing support, preparation of suitable vessels and monitoring
448 technology. https://www.google.com/patents/DE102014005549A1?cl=en (2014).
449 9. Mobley, A., Linder, S. K., Braeuer, R., Ellis, L. M. & Zwelling, L. A survey on data
450 reproducibility in cancer research provides insights into our limited ability to translate
451 findings from the laboratory to the clinic. *PloS one*. **8**, e63221 (2013).
452 10. Ullmann, A. & BÖHM, M. Rfid transponder.
453 https://www.google.com/patents/US20100243742 (2010).
454 11. Ananthanarayanan, V. & Thies, W. Biocoder: A programming language for standardizing
455 and automating biology protocols. *Journal of biological engineering*. **4**, 13 (2010).
456 12. Linshiz, G. et al. PaR-PaR laboratory automation platform. *ACS Synth Biol*. 2, 216-22
457 (2013).
458 13. Soldatova, L. N., Aubrey, W., King, R. D. & Clare, A. The EXACT description of biomedical
459 protocols. *Bioinformatics*. 24, i295-303 (2008).

14. Ellefson, L. P. Label stripping apparatus for label printers. http://www.google.tl/patents/US4036132 (1977).

15. Excoffier, J. L. & Ehrlich, L. E. Sample identification utilizing RFID tags. http://www.google.tl/patents/US20060213964 (2007).

16. Trueeb, H., Birrer, A. & Brauner, T. Method and System to Localise and Identify Test Tubes. https://www.google.com/patents/US20100025464 (2010).

17. Trueeb, H., Birrer, A. & Brauner, T. Laboratory device for processing samples and methods using the same. http://www.google.tl/patents/US8197750 (2012).

18. Pedrazzini, G. System for automatically identifying, conveying and addressing biological material specimens. https://www.google.com/patents/US20100300831 (2010).

19. Barrett, T. J., Check, C. J., Mauch, C. A., Lauria, M. & Scharpf, P. G. RFID label printing system. https://www.google.ms/patents/US6593853 (2003).

20. Hammond, P. A., Ali, D. & Cumming, D. R. A system-on-chip digital pH meter for use in a wireless diagnostic capsule. *IEEE Transactions on Biomedical Engineering*. **52**, 687-94 (2005).

21. Johannessen, E. A. *et al*. Implementation of multichannel sensors for remote biomedical measurements in a microsystems format. *IEEE transactions on bio-medical engineering.* **51**, 525-35 (2004).

22. King, R. D. *et al*. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature*. 427, 247-52 (2004).

23. Soldatova, L. N., Clare, A., Sparkes, A. & King, R. D. An ontology for a Robot Scientist. *Bioinformatics*. **22**, e464-e71 (2006).

24. Kitson, P. J., Rosnes, M. H., Sans, V., Dragone, V. & Cronin, L. Configurable 3D-Printed millifluidic and microfluidic 'lab on a chip' reactionware devices. *Lab on a Chip*. **12**, 3267-71 (2012).

25. Wu, G., Bashir-Bello, N. & Freeland, S. J. The synthetic gene designer: a flexible web platform to explore sequence manipulation for heterologous expression. *Protein expression and purification*. **47**, 441-5 (2006).

26. Vehniaeinen, A., Gustafsson, H. & Koskinen, T. Method and a system for producing nanocellulose, and nanocellulose. https://www.google.ch/patents/US20130303749 (2011).

27. Das, R. *et al*. Integration of photosynthetic protein molecular complexes in solid-state electronic devices. *Nano Letters.* **4**, 1079-83 (2004).

Acknowledgements

Competing interests

The authors declare no competing financial interests.